

Human chromosome 22 and the virtues of collaboration

Two rival approaches to gene sequencing have demonstrated their complementarity with the fruitfly. For similar progress with the human genome, hostilities over data release policies should be re-examined.

Geneticists and many others have cause to celebrate this week. The consortium behind the publicly funded international Human Genome Project has just put the one-billionth base pair of sequence into the public databases, and today is publishing the first sequence of a human chromosome, number 22 (see pages 447, 467 and 489). Given the scale and complexity of the task, the sequence is a great achievement.

The consortium's major rival, Celera Genomics, has itself also just completed the 180 million base-pair genome of *Drosophila melanogaster* (see *Nature* 401, 729; 1999). It did this in collaboration with an academic team headed by Gerry Rubin at the University of California at Berkeley, funded by the US National Institutes of Health (NIH). Participants in the collaboration saw it as a precursor for similar cooperation on the human genome. Their agreement stated that it was "also a test of whether a public/private collaboration can expedite and lower the cost of the generation and use of genomic information".

"Yes" seems to be the answer. Rubin's group provided Celera with maps of the fly genome, for example, and Celera quickly produced additional high-quality sequence data. It has been a win-win affair.

Tensions have so far blocked public/private collaboration on the human genome. Talks between Celera and the US Department of Energy foundered late last year on the thorny question of data access (see *Nature* 397, 93; 1999). The NIH and Britain's Wellcome Trust opposed the deal as conflicting with the so-called 'Bermuda agreement' under which sequence data obtained through the Human Genome Project are released into public databases within 24 hours (see <http://www.sanger.ac.uk/HGP/policy-forum.shtml#ref1>, which includes a concisely readable description of the sequencing process).

Craig Venter, Celera's president, has promised to release human

sequence data quarterly — his subscribers have immediate access. The company argues that the Bermuda agreement is less relevant than it was. What Celera is seeking to patent — as indeed are many scientists at NIH and elsewhere — are genes of demonstrable utility, rather than raw sequence data and genes of unknown function, whose patenting the Bermuda agreement was indirectly designed to pre-empt.

Rubin and Celera managed to strike a pragmatic compromise that included Celera depositing data in GenBank on completion of sequencing. But the value of the cooperation rested less on simply seeing data, and more on the fact that scientists from both sides worked together on the strategy for assembling the sequence data and identifying the genes they contain. The result has been faster and better for it.

It is increasingly clear that the different approaches being taken by Celera and the Human Genome Project are complementary. Collaboration could provide increased coverage and accuracy of the human genome and a higher rate of discoveries important for human health; importantly, a comparison of the two sequences would also yield many single-nucleotide polymorphisms, the major form of DNA variation involved in human traits and diseases.

More substantial collaboration would certainly accelerate completion of the human genome. One question is: at what price in future licensing requirements that would impede research? No progress has been achieved in bridging the public and private genome projects, mainly because the current policies on data release are still far apart. Venter's commitments to his investors are his affair; if he and the Human Genome Project participants can find a mutually acceptable moral and expedient policy on data release, the significant benefits of closer cooperation between Celera and the publicly funded project would be realized. That possibility should be pursued. ■

In pursuit of broadband visions

Ambitious research agendas should stimulate vigorous demand for investment in broadband networks.

As hardware prices continue to tumble, Internet use in countries with low-cost access to the Internet, particularly Scandinavia, Canada and the United States, has penetrated as many as 50 per cent of households. But it is still the science research community that sets the technical challenges, with research centres in neuro-imaging and particle physics, for example, expecting to generate 10^{15} bytes of data every year, and with the anticipated development of petaflop (10^{15} floating-point operations per second) computing power distributed over networks.

The US research community has already set the infrastructure agenda by promoting the idea of 'grids', in which reliable processing power and data are distributed among relevant communities by means of 'middleware' (see, for example, www.gridforum.org). Commendably, the Academia Europaea and the European Science Foundation are also pushing for urgent development along these lines, and are preparing a joint report that urges European action at every level.

So who needs to do what in Europe? A strong report — well documented and vivid in depicting the opportunities — needs to be published as soon as possible, and circulated to all capable of actively supporting the idea, whether in or beyond the research community. A proposal for investment needs to be considered by the council of research ministers next May, and the European Commission should approve funding as part of its fifth Framework programme.

Because such development will also benefit people beyond the research community, the case is intrinsically strong at these elevated levels of decision-making. But researchers must make a case for such investment locally if their institutions are to exploit the grids. The Academia Europaea/European Science Foundation have identified the cost of periodic upgrades that institutions need to provide for: about 10 per cent of their annual budget every seven years or so. That represents a tough challenge for a university. Researchers in the United States, Europe and elsewhere should lobby their rectors accordingly. ■

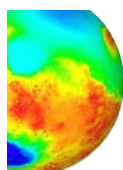




Catch 22
What chromosome sequencing tells us, and which come next
p448



Laser inquiry
Mismanagement blamed for delays at US laser project
p449



Seas on Mars?
Spacecraft sends contradictory data on Martian 'ocean'
p452



Cabinet post
India's top science adviser is champion of the bomb
p453

'Finishing' success marks major genome sequencing milestone ...

Paris

Genome scientists in Britain and elsewhere have shown that they can surmount the enormous difficulties of closing the gaps between ordered chunks of sequence in the human genome — the 'finishing' step.

The evidence comes with the complete sequence of chromosome 22, published in this issue (see pages 467 and 489), and confirms the achievement as a milestone for the publicly funded international human genome consortium.

Finishing is expected to take as much time and effort as the raw sequencing of the 3.5 billion base pairs of the human genome itself, and the completed chromosome 22 gives a foretaste of what the genome project will yield by around 2003, when all the chromosomes are completed (see over).

In contrast, the draft of the human genome promised for next spring by the international Human Genome Project (HGP) will be only a rough line-up of ordered DNA stretches peppered with gaps.

"The significance [of this week's publication] is that it shows you can get very good finishing using the clone-by-clone approach," says Ian Dunham, head of the



Population of geneticists: Dunham (inset) heads the team that sequenced most of chromosome 22.

international team that sequenced chromosome 22. Dunham works at Britain's Sanger Centre, which did most of the work along with eight other laboratories.

"It is the largest contiguous sequence of DNA, and a major breakthrough," adds Michael Morgan, chief executive of the Wellcome Trust's genome campus, one of the main sponsors of the HGP.

In the consortium's clone-by-clone

sequencing strategy, each DNA fragment is cloned and propagated in a library, by inserting it into the genome of a bacterial artificial chromosome (BAC). The jumbled set of individual BAC clones is then rearranged into a 'physical map', typically by looking for overlapping fragments sharing short sequences of DNA that can be identified using the polymerase chain reaction. These are known as sequence tagged sites.

Once BAC 'contigs' spanning the entire genome have been constructed, the BACs, each about 40,000 to 400,000 base pairs long, are sequenced. The hard part is rearranging the fragments in the order in which they occur on the chromosome.

This finishing involves not just joining up raw DNA data to form contigs — groups of clones representing overlapping regions of the genome — but fixing sequence errors, and filling in gaps between contigs. The latter requires extra data and sophisticated techniques to obtain spanning clones anchored to landmarks either side of the gap.

Several scientists argue that finishing may be the Achilles' heel of Celera Genomics, the private US rival to the HGP, which is pursuing a different genome assembly approach. It is processing the millions of DNA fragments — or reads — directly, with no idea of where they might belong in the genome, and then feeding the results into a supercomputer. Craig Venter, its president,

... as researchers pounce on glut of data

London

"A fantastic resource for looking at human population genetics and disease gene susceptibility in a way that has just not been really possible until now." This is how James Scott, professor of molecular medicine at Hammersmith Hospital in London, describes the availability of the complete sequence of chromosome 22 announced in this issue of *Nature*.

Scott has pioneered the field of messenger RNA editing, and has mapped the genes relevant to his research into cardiovascular disease to chromosome 22.

Benefiting from the early release of sequence data from the public Human Genome Project, Scott has already identified seven new genes at the locus, and is now working on their function and organization.

"We could not have done this work without the chromosome 22 data," says Scott, who welcomes the end to the long delays in searching for genes using positional cloning, which has been made obsolete by the availability of the complete sequence.

Single-nucleotide polymorphisms and linkage disequilibrium studies will make it

possible to study the role of genes in heart disease at the level of individuals and populations, he says. A badly working chromosome 22 is also responsible for the congenital heart disease Di George syndrome, which is caused by the loss of part of the chromosome, 22q11.

Genes linked to schizophrenia, a severe mental illness that affects around one in every 100 people worldwide, also sit somewhere on the chromosome. With the full sequence now openly available, it is unlikely to take long to track them down.

D. B.

► hopes to use sophisticated software to assemble the enormous jigsaw by matching their sequences to work out which fragment goes where.

Venter has won over many sceptics with his recent completion of the 180-million-base-pair genome of the fruitfly *Drosophila melanogaster* (see *Nature* 401, 729; 1999). But Philip Green, a biocomputing expert at the University of Washington, points out that the human genome is not only much larger — about 70 million fragments will need to be assembled — but also contains many more families of repeat sequences.

The main problem in sequencing the human genome is that, much more than in lower organisms, it contains many repeat sequences of DNA that look identical and are therefore difficult to place. While Dunham's team have managed to resolve all but 11 of the gaps in the 'euchromatic' region of chromosome 22, they stopped short of attempting to unravel the repeat minefield of the 'heterochromatic' regions.

These are not only difficult to sequence, but are full of repeats. This makes it difficult to resolve the sequence stretches on the short 'p' arm of chromosome 22 from similar repeated heterochromatic regions on chromosomes 21, 13, 14 and 15, for example.

Nonetheless, the success of the clone-by-clone approach in resolving repeats in the euchromatic region has galvanized its supporters. Dunham attributes much of this success to the modular design of the clone-by-clone approach. Given a problem such as a region whose depth of coverage is too low — or which contains gaps or repeats — the modular strategy allows researchers to pull the clone for that region from the fridge and do more work on it. It also allows sequencing to be customized to specific regions, such as GC-rich zones, he adds.

Green argues that the clone-by-clone approach is superior to Celera's whole-genome approach as it reduces the scale of the problem, since each clone will have fewer copies of any given repeat. Finishing techniques such as chromosome walking are difficult on a single clone, and sceptics predict that they would be intractable at the level of the whole genome.

Celera's approach to finishing lacks this safety net and its effectiveness in the human genome remains largely unknown. Also, Celera's approach means that it will only be able to attempt finishing late in the day, when it has the entire sequence.

"The self-sufficiency of Celera's whole-genome strategy for the human will never be put to the test," asserts David Page, a genome scientist at the Whitehead Institute at the Massachusetts Institute of Technology. "Celera's proposed sequencing of the human genome will fully exploit and be utterly dependent upon publicly available HGP mapping and sequencing data." **Declan Butler**

Tiny chromosome is rich in genes and medical promise

London

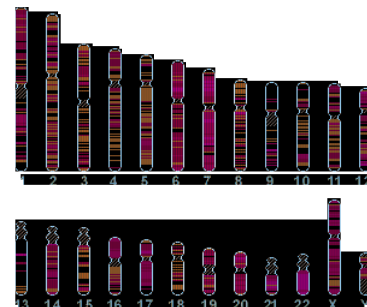
If all the cell's a stage, and the chromosomes merely players, we still don't know whether chromosome 22 has a bit-part or a leading role. But despite its diminutive size — it is one of the smallest of our 23 pairs of chromosomes — we do know it is likely to be quite gene-rich.

So far, for example, chromosome 22 has been implicated in schizophrenia, chronic myeloid leukaemia and trisomy 22, the second most common cause of miscarriages. At least 27 diseases involve some genetic component on the chromosome, which is frequently involved in tumour progression.

About half of its genes, however, have yet to be characterized. Their sheer variety, in terms of differing lengths, is of considerable interest. But the relevance of this won't be known until we have more information from other human chromosomes. So which will be sequenced next?

Greg Schuler, group leader of genome resources group at the National Center for Biotechnology Information, says his short list of candidates would include chromosomes 21 and 7, based on the amount of these sequenced so far (see figure).

"The emphasis on the draft map of the human genome has delayed finishing these chromosomes a little, and researchers are



One down, more to go: sequencing of the other chromosomes is progressing at varying rates.

planning for these next year," says Schuler, adding that the Sanger Centre itself is projecting completion of chromosome 20.

"Things change quite rapidly," Schuler says, so it is difficult to be precise about which will be completed next. The chromosomes are at least partially distributed to different labs, and each of these has its own way of prioritizing.

Completion, says Schuler, is "an accident of how [the labs] decide to queue them up". And what happens when all the chromosomes are complete? "We'll have a big party, of course," he says.

Natasha Loder

Unfinished sequence — the catch on 22

London

Publication of the sequence of chromosome 22 (see pages 467 and 489) means that, for the first time since the Human Genome Project was proposed in the mid-1980s, researchers have an almost continuous sequence for a whole chromosome.

There is, however, a catch. The published sequence remains 'almost continuous' — there are still 11 irritating unsequenced gaps, accounting for about 3 per cent of the sequence. When researchers call the sequence 'complete', they mean they've gone as far as possible with routine methods.

Why do these gaps exist? Researchers on the Human Genome Project sequence DNA by isolating small pieces of chromosome inside bacterial

clones or BACs (see previous page) which multiply to produce sequenceable quantities of the DNA segment. But, for various reasons, clones for certain sequences are simply absent from the collection of clones, or 'clone libraries'.

According to Ian Dunham of Britain's Sanger Centre, the head of the team that sequenced chromosome 22, most of the gaps contain a sequence that is not stable in *Escherichia coli* or yeast. One gap does have a corresponding clone; but, for some unknown reason, researchers are unable to sequence it.

As Peter Little, of Imperial College London, puts it, many pieces of DNA "do really weird things when you try to work with them, and we have no idea why".

Based on what he has seen in chromosome 22, Dunham speculates that missing sequences will tend to occur in certain regions of chromosomes: clustered at the ends, for example, with some gaps at the centres involving repeated sequences.

"The good news is that we can now see how big the gaps are," says Dunham: they are between 50 and 150,000 base pairs. If an interesting gene is thought to be in one such gap, many strategies currently being developed could close these gaps.

"It's a matter of fiddling — there is no reason to suppose these pieces cannot be done," says Little, adding that it would make very little difference to the overall list of genes on this particular chromosome. **N. L.**

US defence labs face research cutbacks ...

San Diego

National defence laboratories in the United States have had to tighten their belts to meet last-minute cuts in funding for long-term research projects chosen at the discretion of their directors. The cuts were quietly imposed by Congress in September as it raced to meet its budget deadline.

The Los Alamos National Laboratory in New Mexico, for example, has introduced a 90-day moratorium on hiring scientists while its officials consider how to reorganize their research in the light of the budget reductions. Some scientists can be hired during this period, however, but only if special recruitment needs arise.

Officials at the Sandia National Laboratories say they have imposed a special review process under which only the most critically needed scientists will be hired for its centres in Livermore, California, and Albuquerque, New Mexico. Sandia expects to receive about \$51 million this year for its discretionary research fund. Last year it received about \$79 million.

The federal budget for the fiscal year 2000, as approved by Congress, calls for discretionary research and development funds to be reduced from six to four per cent of each laboratory's budget. The directors of the laboratories can use these funds, distributed by peer review, to begin promising scientific projects, and the money is considered critical for underpinning weapons development and scientific programmes.

"In the long run, this is a major setback," says Klaus Lackner, acting associate director for strategic and supporting research at the Los Alamos laboratory. According to officials, the lab will receive about \$45 million for the research fund, compared with \$70 million last year. "We're eating the seed corn of our future," he says.

One Los Alamos staff member describes the cuts as "demoralizing" for researchers, claiming that some young researchers have begun looking elsewhere for positions because "they don't see a future" at the lab.

"Weapons are a high-technology business," says Charles Meyers, deputy director for operations and lab development at Sandia. "You have to keep an eye on emerging technologies. The sense of our scientific community is that, if you cut research and development, you are not going [to achieve the desired goals]."

Programme managers at the Lawrence Livermore National Laboratory in California say they are studying how to implement the cuts with the least damage. Top officials at the laboratory, who came under fire last week for mismanaging the construction of the National Ignition Facility (see right), declined to be interviewed.

But they said in a statement that they "are extremely concerned about the impact of this cut" that "will undermine the basic foundations of our current and future programme". Livermore's research fund is being pared to \$35 million from \$58 million, said a spokesman.

The non-weapons laboratories owned by the Department of Energy, such as the Lawrence Berkeley National Laboratory, also in California, spend less on research conducted at the discretion of their directors and will not be affected by the congressionally-imposed ceiling on it. But officials say a 30 per cent cut in travel expenditure imposed across the board by Congress on the Depart-

ment of Energy, labs, could have an impact on the ability of scientists to travel to conferences. "It's not a very constructive environment to do science," says Mike Chartock, head of planning and communication.

Some claim that Congress's cuts in research funds have been triggered by growing concern over management failures by the energy department. Others blame the cuts on political fighting between the Clinton administration and the Republican-controlled Congress at the beginning of an election year. A statement issued by the energy department argues that the reduction "will significantly impact long-term weapons programme milestones".

Rex Dalton

...as Livermore blasted for laser project delay

San Diego

A review committee set up by the University of California has concluded that mismanagement and poor planning are to blame for significant cost overruns and delays in construction of the National Ignition Facility (NIF), the world's largest laser, at the Lawrence Livermore National Laboratory.

The ten-member committee, headed by Steve Koonin, provost of the California Institute of Technology, was set up by the University of California's President's Council to examine construction problems at NIF that were identified in August (see *Nature* 401, 101; 1999). In a 12-page report issued last week, it sharply criticized managers at the Livermore laboratory — which is run by the university on behalf of the US Department of Energy (DoE) — as well as the department itself.

"The project faces serious issues because the DoE and the laboratory violated some basic principles of sound project management," says the panel. "Neither the laboratory, the university nor the DOE had an effective and critical project review process in place."

NIF, the centrepiece of US efforts to ensure the continued viability of its nuclear weapons in the absence of a testing programme was initially to have cost \$1.2 billion and be completed



Mismanaged? NIF construction has been \$300 million over budget.

in 2003. But cost overruns of \$300 million are now expected, with completion delayed by at least 18 months.

The review committee endorsed the technical feasibility of the project, which has 192 lasers for use in nuclear, fusion and/or scientific experiments. But construction management was frequently deficient — project planning significantly underestimated contingencies, the committee said, and baseline goals were set too early.

The laboratory also had an ill-conceived "do-it-yourself mentality" that discouraged the use of outside expertise, the committee says, and "there were multiple failures at multiple levels that kept appropriate people ignorant of these concerns". The laboratory's director needs to "take ownership" of management of the project, it adds.

DoE secretary Bill Richardson

had told lab employees and the media gathered for the unveiling of the NIF target chamber last June that the project was on budget and on time. Even then, however, scientists at the laboratory and officials at the DoE knew there were serious problems. Officials now say they failed to brief Richardson adequately before his speech.

Bruce Tarter, Livermore's director, last week admitted past "weaknesses in the project's management structure", and said that significant steps — including replacing management, reorganizing teams and rededicating lab personnel — have already been taken to address the issues identified by the committee and internal reviews.

The DoE said in a statement that the review provided evidence that the university was committed to getting the NIF project back on track.

R. D.

EMBL rescue package keeps bioinformatics centre running

Munich

The threat to the European Bioinformatics Institute (EBI) arising from the European Commission's refusal to fund its running costs seems to have been temporarily lifted.

At a meeting of the council of the European Molecular Biology Laboratory (EMBL) in Heidelberg last week, delegates of EMBL's 16 member states agreed in principle to make up the shortfall of 5 million euros (US\$5 million) for next year. EBI, based near Cambridge in the United Kingdom, is an outstation of EMBL.

Although EBI scientists are relieved, they have not relaxed completely. Some member states were unable to pledge money immediately, as their research budgets for next year are already fixed, and they will have to make the case for funds to be redistributed. A final decision awaits a special meeting of the EMBL council in March.

Britain will advance money if needed for the first three months of next year to keep EBI fully functioning. Britain is by far the biggest user of the institute, with the number of accesses each month averaging almost twice that of the next most frequent users — the United States and Germany. It is also a strong supporter of EBI's further development.

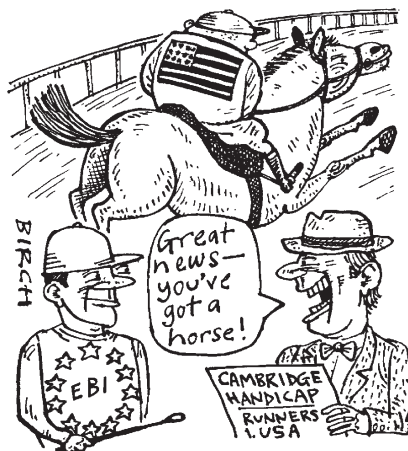
"This is a vote of confidence for EBI," says Graham Cameron, the institute's joint director. "But we are not out of the woods yet." Although relieved that immediate rescue funds have been found, Cameron stresses a widely held view that "EBI's budget needs to double if it is to become a competitive global force in bioinformatics".

The National Center for Biotechnology Information, the institute's US equivalent, has well over twice its budget. Cameron fears that bioinformaticists, in heavy demand from industry as well as academia, will leave EBI if its prospects do not brighten soon.

Julio Celis, chairman of the EMBL council and head of the Danish Centre for Human Genome Research in Aarhus, shares this concern. "The council meeting last week recognized the importance of bioinformatics as central to many areas of biological and medical research in the next millennium".

He points out that, while confident that member states will come up with the money at the March meeting, "this will only restore EBI's budget to its current level". The council, he adds, will have to be prepared to step up a gear in financing the next EMBL research plan, which starts in 2001.

EBI's success is reflected in its number of accesses, which is rising by 20 per cent per month. The institute expects this to reach 25 per cent, and stay there, by the time the first



draft of the human genome sequence is completed in the spring.

Bioinformatics is a huge growth area, and Europe must handle the fact that an entirely different level of stable funding must be made available, says Celis.

The new research commissioner, Philippe Busquin, is sympathetic to this view, but says that there is no way around the rules of the European Union's fifth Framework Programme (FP5). These allow funding for research projects, but not for the facilities themselves, although some scientists argue that the rules can be interpreted as allowing core support (see *Nature* 402, 3 & 4; 1999).

Busquin believes that both EBI and the European Mouse Mutant Archive (EMMA), a facility based near Rome that has also been thrown into crisis by withdrawal of European funds, must be funded at a European level. But neither can wait until new rules are designed for FP6, which should start in 2003.

At an informal meeting of the council of European research ministers, Busquin will launch a debate on infrastructure funding. He hopes to prompt member states to draw up agreements to finance EBI, EMMA and similar facilities, and to build a consensus for infrastructure funding in FP6.

A spokesman for the research commission points out that member states found a solution for EBI, "though not an ideal one". Member states that voted against core funding of all infrastructure in FP5 may be more discriminating in FP6, he notes.

"It is somewhat ironic that there is so much unallocated money in the FP5 infrastructure pot yet, to solve the EBI crisis, EMBL has to persuade 16 different governments to divert money earmarked for other national purposes, at short notice," says Glauco Tocchini Valentini, director of the CNR Institute of Cell Biology at Monterotondo and a member of EMBL council.

Alison Abbott

Centralized German research centres will compete for funds

Munich

Germany's national research centres have agreed to abandon their independence and accept a plan to centralize decision-making on strategic research priorities, a move that will increase both competition and collaboration between them.

The plan was announced last week at the annual assembly of the Hermann von Helmholtz Association (HGF), the umbrella group for the 16 centres. Of their DM3 billion (US\$1.6 billion) annual budget, 90 per cent is provided by the federal government, with the remainder coming from *Länder* (state) governments.

Six strategic 'cornerstones' have been chosen for the institutes' research programmes: the structure of matter, environmental and Earth sciences, transport and space research, health, energy, and enabling technologies.

Each cornerstone is subdivided into several areas. Health, for example, covers biomedical research, medical technologies and public health research. Biomedical research includes areas such as cell-cell and virus-cell interaction, functional genome and proteome analysis by DNA and protein chip technology, gene therapy and bioinformatics.

The first national research centres were set up in the 1950s to carry out nuclear research. Changing political and technological priorities mean that their focus has shifted to medical and environmental research, but recently their research activities have been criticized for a lack of scientific coherence.

The plan will be overseen by a commission of representatives of the institutes and of the federal and regional governments. Under the proposed reorganization, the HGF's senate would act as a supervisory and decision-making board with broad responsibilities for determining strategic research priorities, and allocating money accordingly.

The influential board will comprise members of the HGF, the federal government and the *Länder*, and also representatives from industry and the heads of Germany's major research organizations — the Max Planck Society, the Fraunhofer Society and the Conference of University Rectors.

The centres would report to — and be financed by — a body responsible for the general direction of their research, and will have to compete for research funds. All proposals will be reviewed by panels of external experts chosen by the senate.

► The groundwork for the shift was laid two years ago, when the centres agreed to divert five per cent of their budgets to a central, DM150-million strategy fund, to which they have to apply competitively (see *Nature* 387, 837; 1997). "The strategy fund has prepared the ground for more collaboration and competition between the centres," says Detlev Ganten, scientific director of the Max Delbrück Centre for Molecular Medicine in Berlin and head of the HGF. "But now we will have a de facto global budget."

Most HGF members welcome the changes, hoping that they will sharpen the centres' scientific profiles at a time when the Wissenschaftsrat, Germany's science council, has begun a systematic evaluation of the HGF. "Provided that basic research is not excessively damaged, it is certainly the right development," says Hans-Günter Afting, director of the National Research Centre for Environment and Health in Munich.

But some of the HGF's research may have to find a new home. The Bonn-based National Research Centre for Information Technology, for example, will leave the HGF in 2001 and join the Fraunhofer Society, Germany's organization for applied research. According to Ganten, a further reduction of the number of national research centres is "within the bounds of possibility".

Quirin Schiermeier

Nuclear safety and biotech boosted in Japan's budget

Tokyo

Biotechnology, venture businesses and nuclear safety all receive extra support in Japan's latest supplementary budget, approved by the parliament last week, which is aimed at putting the nation's economy on a full recovery track.

As well as a strong emphasis on public works projects, the supplementary budget — the second this year — includes generous support for science. This is largely a result of prime minister Keizo Obuchi's Millennium Project, which seeks to promote new industries in areas such as biotechnology, information sciences and environmental technology (see *Nature* 401, 313; 1999).

The supplementary budget totals ¥6.79 trillion (US\$66 billion), of which ¥907.6 billion is earmarked for science and technology, and ¥773 billion for supporting small and medium-sized businesses.

The Science and Technology Agency (STA) will receive ¥16.4 billion for life sciences, including ¥4.3 billion for nuclear magnetic resonance facilities being built at the new headquarters of the Genomic Sciences Centre, ¥1.2 billion for proteomics research, and ¥2.4 billion for the analysis of human complementary DNA (cDNA). The

agency will also receive ¥46.1 billion for additional efforts in nuclear safety, following the accident at a uranium-processing plant in Tokaimura, Ibaraki Prefecture.

The Ministry of International Trade and Industry (MITI), which takes over some of the nuclear safety responsibilities of the STA in 2001, and the Ministry of Health and Welfare will receive ¥30.8 billion and ¥11.6 billion, respectively, for nuclear safety research.

MITI will receive ¥12.5 billion for biotechnology-related research, including proteomics projects, while ¥35.6 billion will support venture businesses, particularly entrepreneurial activities at universities.

A budget boost for the Ministry of Agriculture, Forestry and Fisheries will speed up the rice genome sequencing project by almost four years (see *Nature* 401, 102; 1999).

Public works projects include building at national universities, such as the new campus at Kyoto University. They also include plans for new life-sciences-related research institutes.

The budget brings Japan's science spending beyond ¥17 trillion, as promised by the Science and Technology Basic Plan, which pledged to double science spending between 1996 and 2000.

Asako Saegusa

Wellcome Trust backs Rutherford to host synchrotron

London

Britain's Wellcome Trust has declared its preference that a new Anglo-French synchrotron source should be built at the Rutherford Appleton Laboratory (RAL) in Oxfordshire. The trust is providing a substantial proportion of the construction costs.

But the statement coincides with an announcement by the UK government that the long-awaited decision over a home for the facility, known as Diamond, will be delayed again to allow two new reports on the candidate sites to be commissioned.

Trade and industry secretary Stephen Byers also said last week that he has asked John Taylor, the director-general of research councils, to seek the views of synchrotron users "on what they see as the key issues of relevance to a decision on the site".

An engineering survey and report are also being commissioned from consultants for the two potential sites — at RAL and at Daresbury Laboratory near Manchester. A decision is now expected in mid-January.

Science minister Lord Sainsbury acknowledged in a letter to the *Financial*



Sainsbury: tough questions over science synergy.

Times newspaper this week that there were "difficult scientific considerations to resolve" about the synergy between the synchrotron and the existing neutron spallation source at RAL, "as well as issues concerning the suitability of the two sites".

In its statement last week, the Wellcome Trust says it prefers RAL because it believes that greater scientific benefits would result from a location close to the existing neutron source and to Medical Research Council units and the University of Oxford.

But union officials and some senior synchrotron scientists say this factor is a "small but not overriding consideration", given the small numbers of scientists that would benefit. The trust has also generated anger among supporters of Daresbury by suggesting there are greater planning risks associated with the northern site.

Chris Brough, assistant director of plan-

ning for the area, criticizes "unsubstantiated assertions" by the trust's planning consultants about potential drawbacks of the Daresbury site. He says they have failed to do the "most basic thing that any planning consultant worth its salt would do — contact the local planning authority".

"It's foolish to say we are high risk [when] we have a sympathetic development agency," says Sue Smith, a union representative at Daresbury. "RAL does not even have potential [outline] planning permission."

The dispute over where Diamond should be sited has raised fears that choosing the Rutherford site would eventually lead to the closure of Daresbury and further deepen Britain's north-south divide.

Daresbury appears to have strong support from local government, local Members of Parliament, and Byers. The unions say they have a letter from an Oxfordshire councillor expressing concern at the prospect of economic "overheating" in the area around Rutherford and at problems of providing housing for the existing influx of workers into the region.

Natasha Loder

Opinion divided on images of Martian 'ocean'

Washington

Researchers remain split on whether a large ocean — or oceans — existed on Mars some two billion years ago, when the now desert-like planet was warmer and wetter.

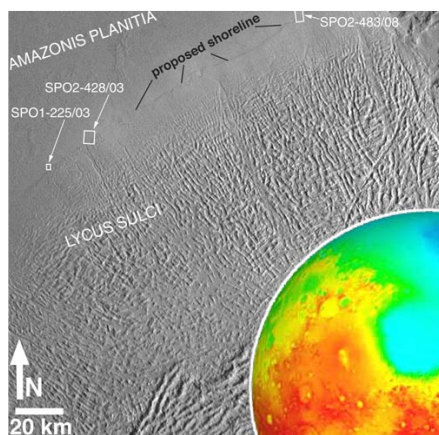
The theory has intrigued planetary scientists for a decade, ever since Timothy Parker of NASA's Jet Propulsion Laboratory discerned in Viking spacecraft images what looked like ancient coastlines bordering dried-up basins in the planet's northern lowlands.

Recent results from the Mars Global Surveyor (MGS), which is currently orbiting the planet, have provided evidence both for and against the ocean interpretation — but nothing approaching a verdict.

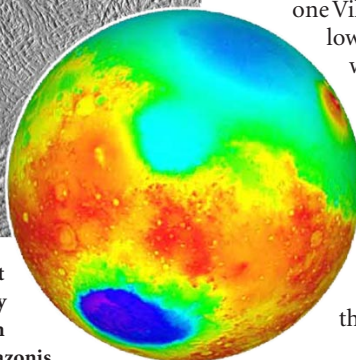
The Mars Orbiter Laser Altimeter (MOLA) on MGS, which measures topographic relief on the surface, has shown the northern lowlands to be even flatter than previously thought — as flat, in fact, as the abyssal plains beneath Earth's oceans.

While this alone proves nothing (lava flows can also be flat), the MOLA results were what might have been expected from an ocean floor. Estimates of the volume of water this basin could have held also fit neatly with estimates of the volume drained by channels still visible on the surface and the amount of water the subsurface could hold after the ocean disappeared.

Using MOLA data, James Head of Brown University, in Providence, Rhode Island, has checked different locations along the lower of Parker's two 'shorelines' ('Contact 2') and



Blue planet? The jury is out as to whether the unusually flat regions in the northern lowlands of Mars (the 'Amazonis Planitia' in the photo above, and blue at upper right) are evidence of an ancient ocean.



found them to be very nearly the same height, supporting the notion of a level ocean rim. The boundaries of this proposed sea even correlate well with a geologic unit known as Vastitas Borealis (the 'northern vastness').

But when a team at Malin Space Science Systems in San Diego examined more than a dozen high-resolution views of different sections of the purported shoreline taken by the MGS Mars Orbiter Camera, they came up empty-handed.

Michael Malin and Kenneth Edgett reported in the 1 October issue of *Geophysical Research Letters* that "none of the images that cross the proposed coasts exhibit features that could be interpreted as having an obvious or unambiguous littoral origin".

What looked like wave-cut cliffs in one Viking picture turn out to be shallow slopes. In another place, where Parker and others had drawn a boundary line for the hypothetical ocean, Malin and Edgett found nothing. But even they say that such a meagre sample cannot settle the matter. By Head's reckoning, Malin and Edgett have examined only 0.04 per cent of the mapped Contact 2.

Parker admits that the high-resolution views have been "very frustrating" so far, but has found other images from the same Mars Orbiter Camera that support the ocean hypothesis. He is combining MOLA elevation data with Viking imagery to refine his maps, while continuing to study the close-up images.

Part of the problem, according to Head, is knowing what to look for. Would Martian coastal features have been sculpted by crashing ocean waves, or by subtler processes akin to those found in terrestrial lakes? Would the surface have been liquid or ice?

The tug of gravity is different on Mars, and there is no large moon to generate strong tides. Any shoreline would probably have experienced tectonic change and two billion years of wind action. Sorting all this out, says Head, is a "long, tedious, arduous task", unlikely to yield quick answers.

Other researchers are adding pieces to the puzzle. James Garvin, of NASA's Goddard Spaceflight Center, has found that crater morphologies inside the 'ocean' basin are strikingly different from those outside it. The small, shallow craters inside the basin could be explained by ground ice or loose sediments at the impact site.

The Thermal Emission Spectrometer on MGS, which maps surface mineralogy, has found a general lack of carbonates and evaporites, not just in the northern lowlands but all over the planet. If there was an ocean on Mars, its chemistry may well have been different from the seas on Earth.

Edgett, who is unconvinced by the high-resolution pictures he has examined so far — no clear coastal landforms have emerged from the images he has studied since the October paper was published — claims to be agnostic on the subject of the Mars ocean.

But one thing is certain. Martian geology is turning out to be strange and complex. "It's really alien," he says.

Tony Reichhardt

South African telescope to go ahead

Cape Town

Ben Ngubane, South Africa's minister of arts, culture, science and technology, gave the green light last week for the construction of the Southern African Large Telescope (SALT), the largest single telescope for optical and infrared astronomy in the Southern Hemisphere.

The design of the new telescope will be based on that of the recently completed Hobby-Eberly Telescope at McDonald Observatory, in Texas, United States. The decision to authorize construction followed the signing of an agreement on science and technology between Poland and South Africa by Ngubane and his Polish counterpart, Andrzej Wiszniewski.

Poland's contribution to SALT is more than 15 million rand (US\$2.5 million), making up one third of the 45 million rand committed so far by international partners towards the telescope. The South African government has

committed 50 million rand, enabling construction of the telescope to commence next year.

"SALT will enable South Africa to remain internationally competitive in astronomy well into the twenty-first century," says Ngubane. He went on to say that the telescope's activities will range from observing some of the most distant galaxies in the Universe and studying the early evolution of the cosmos to searching for planets around neighbouring stars.

Now that formal approval has been given by the government, the project team can be recruited and the construction of the telescope, expected to take five years, can be started next year.

The telescope will be sited at Sutherland in the Northern Cape, an outstation of the South African Astronomical Observatory operated by the National Research Foundation.

Michael Cherry

France cuts its 'big science' spend to bolster lab research

Paris

French research agencies have been told by research minister Claude Allègre to cut FF200 million (US\$31 million) from their budgets for large scientific facilities, to provide extra money for laboratory research.

The national space agency, CNES, which accounts for half of the spending on large programmes, will absorb half of the cuts. To achieve this, the agency will put on hold plans to launch Corot, a space telescope designed to study the seismology of stars and look for extrasolar planets, in 2003.

The Centre National de la Recherche Scientifique (CNRS) and the atomic energy commission, CEA, will share the remaining FF100 million in cuts.

At the CEA, the cuts will affect France's contribution to two particle detectors for the Large Hadron Collider at the European Laboratory for Particle Physics (CERN), in Geneva, Switzerland, and the heavy-ion accelerator at Ganil in Caen.

France's contribution to two satellites planned by the European Space Agency — First, which observes primordial galaxies, and Integral, which studies stellar explosions — will also be affected. The research minister, a long-standing critic of manned space flight, has also said he will cut back on France's contribution to the International Space Station.

On arriving in office two years ago, Allègre announced his intention to cut the

budget for large equipment, which accounts for FF4.6 billion this year and FF4.8 billion next year.

Even though the 8.4 per cent of France's research budget spent on large equipment is no more than that spent in the United Kingdom, Germany and the United States, Allègre considers the spending of these countries excessive.

According to ministry figures, spending on big science has grown by 69 per cent over the past ten years. In contrast, general financial stringency has meant that spending on research projects has increased by 34 per cent over this period, while the overall research budget was up by only 22 per cent.

Allègre's bid to reduce spending on large equipment has prompted him to seek international cooperation for future projects. One result has been the controversial decision in August to join up with British plans to build a new synchrotron, rather than construct a French machine (see page 451 in this issue and *Nature* 400, 489; 1999).

While many researchers applaud Allègre's commitment to laboratory spending, not all are pleased with seeing the money drained from big-science facilities.

"It's true that the CNRS labs are poor," says Edouard Brezin, president of the board of the CNRS. "But these cuts are going to make it pretty difficult to run the machines that we are using. We don't know right now how to resolve this."

Heather McCabe



Kalam: unprecedented role for science adviser.

Champion of India's bomb rewarded with cabinet position

New Delhi

A. P. J. Abdul Kalam, a pioneer in India's rocket and missile programme who was largely responsible for the country's nuclear tests last year (see *Nature* 393, 197–198; 1998), has been appointed principal science adviser to the government. He will be the first working scientist to achieve cabinet rank.

Kalam is currently scientific adviser to the defence minister and chief of the Defence Research Development Organization, which is responsible for about 40 labs. He will now oversee much of India's scientific activity — including atomic energy and space — and advise the government on all aspects of science policy.

Officials in the office of prime minister Atal Behari Vajpayee say that the move, announced last week, is the first step towards restructuring and coordinating India's scientific agencies (known as departments), which have been growing increasingly independent, to ensure they address national priorities.

Kalam is expected to break down barriers among departments and ministries and build synergy into the scientific system. His cabinet rank is intended to help him achieve this.

Scientists have mostly welcomed the appointment. "If the idea of having a principal science adviser is to bring cohesiveness, it is a good move because Kalam can do it," says Ragunath Mashelkar, secretary of the Department of Scientific and Industrial Research.

But Kalam's cabinet rank is expected to create some confusion. Vajpayee's council of ministers already has two science ministers — one being Murli Manohar Joshi, a former physicist. Kalam's new status has also made the job of the junior science minister, Bachi Singh Rawat, redundant.

K. S. Jayaraman

EMBL pay settlement will cost millions

Munich

A long-running labour dispute at the European Molecular Biology Laboratory (EMBL) in Heidelberg, Germany, was settled last week at a cost of at least DM4 million (US\$2 million) to the organization's 16 member states.

EMBL's council agreed to implement a judgement of the Geneva-based International Labour Organization (ILO), which had ruled that the lab was wrong to abandon in the mid-1990s its adherence to the salary recommendations of the Coordinated Organizations, a committee that sets salary scales for international organizations (see *Nature* 375, 442; 1995).

It remains unclear how seriously this compliance will affect EMBL. The council has asked for clarification of whether the ruling refers simply to the salary adjustment from 1995 — which would cost 2.1 per cent

of the laboratory's budget in back pay and DM0.8 million per annum in future — or also to a backdated implementation of higher salary scales. This second option would cost considerably more — namely 8 per cent of the budget in back pay and DM3.5 million per annum.

Member states have agreed to find the money to support the first but not the second scenario, which would have to be financed from within the existing budget. Final decisions will be confirmed at a council meeting in March.

The settlement comes at a time when EMBL already faces the withdrawal of European Commission support for its outstation, the European Bioinformatics Institute (see page 450), and is expecting to seek from member states a substantial increase in its basic budget for its next five-year indicative scheme, beginning in 2001.

Alison Abbott

South Africa's health minister downplays AIDS drug concerns

Cape Town South Africa's health minister, Manto Tshabalala-Msimang, has qualified her support for criticism of the low-cost anti-retroviral drug nevirapine contained in a recent statement from the health ministers of southern African nations. The statement expressed concern that the virus could develop resistance to the drug (see *Nature* **402**, 332; 1999). In a speech to the South African parliament, Tshabalala-Msimang conceded that administering a single dose would be unlikely to cause this, although she admitted that "scientists cannot guarantee that this will not happen".

Tshabalala-Msimang described nevirapine as "safer, less expensive and more practical than AZT or any other drug tested so far in preventing mother-to-child transmission of HIV". She said that the results of studies being carried out in South Africa, to investigate the comparative cost-effectiveness of nevirapine, AZT and AZT plus 3TC, would be available next March.

Lynne Boschoff, medical director in South Africa of Boehringer Mannheim, which manufactures nevirapine, confirmed that "there is no indication that resistance should

be expected as a result of a single dose, provided the patient has not been exposed to the drug before".

Third World academy urged to appoint expert panels

Dakar, Senegal The Third World Academy of Sciences, which brings together the scientific academies of developing nations, has been urged to set up international panels of experts in critical areas such as information technology, biotechnology and biodiversity to advise governments in Africa and throughout the developing world.

The call came in a joint declaration issued at the end of the academy's seventh general conference held in the Senegalese capital of Dakar last week. Attended by 300 individuals from 60 countries, the meeting also urged the academy to draw up a white paper exploring the opportunities and risks of creating partnerships between scientists and the private sector in the developing world.

Canadian award will back top researchers

Washington The Canadian government has announced an annual Can\$1 million award for scientific or engineering research of exceptional quality. The award will be named

the Gerhard Herzberg Canada Gold Medal in honour of the late chemist who won Canada's first Nobel Prize in Chemistry.

The winner will receive Can\$200,000 a year over five years in research funding from Canada's Natural Sciences and Engineering Research Council. "Our intent is not just to celebrate Canadian heroes, but to provide them with a new level of research support that Canadian researchers have never enjoyed before," says Tom Brzustowski, president of the council. The winner of the first award will be selected next year.

Van Duinen to lead Europe's science foundation

Munich Reinder van Duinen, president of the Netherlands' National Research Council (NWO), will become the president of the European Science Foundation (ESF) in January. Van Duinen, 59, a physicist by background, was unanimously chosen last week by the ESF's electoral committee.

He takes over from Sir Dai Rees, the former chief executive of Britain's Medical Research Council. ESF officials hope that Rees, who spearheaded the foundation's activities in the area of risk assessment and society, will maintain his close links with the organization.

This should not be the end for terminator technology in GM crops

Sir—Monsanto's decision to abandon 'terminator' technology has been welcomed by those concerned about the genetic modification of food crops. The public perception is that this technology would force farmers in the developing world to purchase expensive seed year after year from multinational conglomerates. Although this is an important consideration, such technology might also have beneficial consequences, because ensuring the sterility of genetically modified (GM) crops could offer an important mechanism for environmental containment. By genetically engineering sterility, one could ensure that GM plants could not spread or cross-pollinate with naturally occurring varieties.

The paradigm of sterility/incompetence is not new. The laboratory molecular biologist is all too familiar with containment, both at the physical and the biological level. In the 1970s there were fears that genetic engineering would result in the release of 'horror pathogens'. These fears were countered by the engineering of replication-incompetent vectors and disabled host organisms unable to grow outside the lab. This principle has been extended to the situation where molecular biology is used to treat human disease—gene therapy—where ensuring replication incompetence of viruses has been an important regulatory hurdle for virally mediated gene transfer.

Concerns about cross-fertilization have featured frequently in the UK media, and have been countered mainly by arguments about physical distance, such as how far bees can fly. It is surprising that the lessons learned in applying molecular techniques to contain GM organisms in the lab have not been applied to agricultural biotechnology. Such features should be placed high on the agenda of policy-makers and those in industry trying to persuade us of the safety of their products. Although there are moral and social dilemmas surrounding 'terminator' technology, these techniques should not be discarded without further reflection.

Andrew Jackson, Chris Inglehearn

Molecular Medicine Unit, University of Leeds,
St James's University Hospital, Leeds LS9 7TF, UK

Weighing up the threat to the world's corals

Sir—A. J. van Loon's Correspondence, "Corals resist extinction by global warming", shows how two correct, but unrelated and incomplete, statements can

lead to false conclusions¹. Missing pieces in his reasoning include the different genetic base of corals past and present, the amount of temperature change to which corals can adapt, and the role of human-caused stresses.

The corals living during warmer conditions before the Pliocene–Pleistocene transition were different species from those alive today. The start of the Ice Ages saw a period of species extinction and origination². Today's corals are descendants of the most cold-tolerant corals that survived glacial episodes, while corals adapted to high temperatures largely vanished.

The upper temperature limits of modern corals have been determined experimentally by many researchers, starting with Mayor³ and Yonge and Nicholls⁴. They showed that only a few degrees of warming above normal maximum temperatures bleached and killed most corals, and found no behavioural or physiological capacity to adapt to higher temperatures. Subsequent research has refined but not changed their key conclusions^{5–9}.

Extinct species were once able to adapt to warmer conditions than extant corals can tolerate, and presumably modern corals could do so again if there were sufficient evolutionary time and sufficiently slow rates of change. But fossil corals evolved in a world without human-accelerated change. Coral temperature tolerance thresholds are now being exceeded worldwide, and rates of temperature change currently under way dwarf 'natural' rates.

It is hardly surprising that modern corals are incapable of rapidly re-evolving the tolerance of high temperatures that their ancestors appear to have lost millions of years ago. Exponentially increasing anthropogenic pollution—in the form of greenhouse gases, nutrients, pathogens, sediments and other stresses—makes it very unlikely that corals will be able to adapt to a 'greenhouse world' quickly enough to survive, despite wishful thinking in *Nature*.

Thomas J. Goreau

Global Coral Reef Alliance, 324 Bedford Road,
Chappaqua, New York 10514, USA

Sir—Van Loon¹ argues that, just because coral reefs survived the geological past, global change is not a threat to them. It surely depends on the time-scale over which such change occurs. Corals survived sea-level fluctuations, in particular rapid drowning at the last glacial termination, but there are indications that reefs took 1,000 years to become re-established on the new shelves¹⁰.

Today's reef bleaching can be compared to what happened at the Cretaceous/Tertiary boundary, when carbon dioxide

and temperature levels rose shortly after the asteroid impact. Spectacular shell abnormalities are present in reef foraminifers today¹¹, and similar effects occurred just after the boundary¹². Reefs took about 10 million years to recover, mainly by recolonization by small corals or deep-sea forms¹³.

Reef decay will probably not be so great in the future, but the geological record gives a taste of what may be to come.

M. Pecheux

IUFM, Nice University,

15 rue des Roses, 06000 Nice, France

1. van Loon, A. J. *Nature* **400**, 708 (1999).
2. Budd, A. F. & Johnson, K. G. *Proc. 8th Int. Coral Reef Symp.* **1**, 423–428 (1997).
3. Mayor, A. G. *Carnegie Inst. Wash. Pap. Mar. Biol.* **12**, 175–178 (1918).
4. Yonge, C. M. & Nicholls, A. G. *Great Barrier Reef Expedition Sci. Pap.* **1**, 135–211 (1931).
5. Jokiel, P. L. & Coles, S. I. *Coral Reefs* **8**, 155–162 (1990).
6. Williams, E. H. & Bunkley-Williams, L. *Atoll Res. Bull.* **335**, 1–71 (1990).
7. Glynn, P. W. *Trends Ecol. Evol.* **6**, 175–179 (1991).
8. Goreau, T. J. & Hayes, R. L. *Ambio* **23**, 176–180 (1994).
9. Goreau, T. J., McClanahan, T., Hayes, R. & Strong, A. *Conserv. Biol.* (in the press).
10. Davies, P. *et al.* *Proc. 5th Int. Coral Reef Conf.* **3**, 95–103 (1988).
11. Hallock, P. *et al.* *J. For. Res.* **25**, 280–286 (1995).
12. Gerstel, J. *et al.* *Paleoceanography* **1**, 97–117 (1986).
13. IGCP Project 286: *Early Paleogene Benthos* (UNESCO, Paris, 1989).

Who's doing what in US protein initiative

Sir—We read with interest the article describing the initiative by the National Institute of General Medical Sciences (NIGMS) to assemble structural information on at least 10,000 proteins over the next five years (*Nature* **400**, 494; 1999). What this otherwise informative article did not make clear was the role of other US government partners in supporting the basic structural-biology infrastructure upon which this initiative is built—particularly the Offices of Biological and Environmental Research and Basic Energy Sciences of the Department of Energy (DOE), and the National Science Foundation (NSF).

Additionally, the DOE has played a major role in developing structural genomics in the United States. Over the past decade the National Institutes of Health (NIH), NSF and DOE have substantially increased access to synchrotron light sources for protein crystallography through a series of initiatives. The agencies are implementing further improvements at the synchrotrons that will provide the capacity needed to reach the ambitious goals of the Protein Structure Initiative (PSI). The DOE is supporting structural genomics efforts at Brookhaven, Los Alamos and Lawrence Berkeley National Laboratories. In January 1998, the DOE

helped to organize a meeting at Argonne National Laboratory which triggered subsequent meetings at NIH and elsewhere that resulted in the PSI.

All three agencies support the Protein Data Bank, which is managed by the Research Consortium for Structural Bioinformatics to provide access to the more than 10,000 known protein structures. The addition of 10,000 new structures through the PSI will provide the information needed to define the several thousand key protein folds that, in turn, should enable classification of proteins into functional categories. The contributions of each of these agencies (and those of several non-federal organizations) will be needed to enable the NIGMS initiative to succeed.

Marvin Cassman*, **Ari Patrinos†**

*Director, National Institute of General Medical Sciences, National Institutes of Health, Bethesda, Maryland 20892-6200, USA

†Associate Director for Biological and Environmental Research, Department of Energy, 19901 Germantown Road, Germantown, Maryland 20874-1290, USA

Search is on for better search engines

Sir—Steve Lawrence and C. Lee Giles stated in their Commentary that most of the popular search engines index only about 7–16 per cent of the World-Wide Web¹. This is alarming, as many scientific web pages containing important data may never be discovered. As the web grows it is going to become increasingly difficult for general search engines to give comprehensive coverage. The answer to the problem could be the development of subject-specific search engines able to cover most of the contents within that subject.

Most currently available subject-specific lists and indexes are maintained by humans. Many of them are merely collections of web addresses and lack context-based relevance ranking and retrieval of results in multivariate combinations. What is needed are search engines that could traverse through pages at the last level in a subject-specific website. They would be able to do this as the numbers of such sites would be within manageable limits. Crawlers or robots traversing through such a subject-specific web subset could build up a comprehensive and complete bank of keywords. In turn, such keyword-mounted crawlers would efficiently and more frequently screen the last-level page of the site.

Subject-specific search engines would be able to maintain the freshness of the hits, as the crawlers would check a manageable number of specific web pages more

frequently than they would by moving through the entire web.

Mike Gardner² has rightly suggested that we need science-oriented search engines with sets of scientific metadata, as metadata are the key to better searching. Together with that proposal, an approach similar to peer-reviewing of scientific publications could be applied for categorizing and evaluating web pages based on their content, quality and subject-specificity. It would be feasible to use algorithms and rule-based expert systems to check content-richness, subject-specificity and freshness-based context-relevance ranking for retrieved results.

Millions of dollars need to be invested in developing search engines. This investment could be cost-effective if it resulted in an almost zero noise-to-signal ratio and precise but comprehensive subject-relevant hits. The development work should be done in the academic sector, but the completed search engines would have commercial potential, and would generate more revenue than general search engines because of their subject-specificity.

Development of subject-specific search engines would satisfy the growing demand for the latest, precise, value-added, noise-free hits with a high level of subject relevance. Many such search engines together would be able to index most of the World-Wide Web.

Vishwas Chavan

Centre for Cellular and Molecular Biology, Uppal Road, Hyderabad 500 007, India

1. Lawrence, S. & Giles, C. L. *Nature* **400**, 107–109 (1999).

2. Gardner, M. *Nature* **401**, 111 (1999).

Japan builds bridges to rest of the world

Sir—Your Opinion and News article about research in Japan raise an important and timely issue (*Nature* **401**, 309 & 314; 1999). A focus on higher quality rather than quantity of scientific output, and more emphasis on cooperation with researchers abroad, are indeed important goals for Japanese science. But unfortunately the articles did not compare the Japanese situation with that elsewhere.

You report, for example, that 40 per cent of researchers submit papers only to Japanese journals, which is said to reflect insularity. But how can the reader draw such a conclusion without knowing comparable numbers for other countries? I would not be surprised if 40 per cent of researchers in the United States, the United Kingdom or Germany published only in national journals. And you did not mention that a considerable number of Japanese journals publish articles only in

English, which is surely an attempt to reach readers abroad.

It is interesting to learn that 40 per cent of researchers submit papers only to Japanese publications, while 34.1 per cent also submit to international journals. Does that mean that 25.9 per cent only publish in overseas journals? If true, this would show that there are many internationally minded Japanese researchers.

You do a good job of pointing out weaknesses and important goals for Japanese research. But your omission of comparative numbers for elsewhere in the world does not allow readers to see the problem in its true proportions. It could even be the case that your articles inadvertently promote prejudices about Japanese science, which would be unfortunate at a time when many Japanese research teams are striving to attract foreign scientists.

Philippe Buhlmann

Department of Chemistry, The University of Tokyo, Hongo 7-3-1 Bunkyo-ku, Tokyo 113-0033, Japan

Galileo had accurate vision of the Moon

Sir—Martin Kemp writes that “the details of Galileo’s cratered Moons are difficult to align precisely with actual features”¹. I would like to draw readers’ attention to the work of Ewen Whitaker², a selenographer at the University of Arizona. Whitaker reviewed earlier efforts at identifying the features drawn by Galileo, and took special note of the important contributions of Guglielmo Righini, Owen Gingerich and Stillman Drake. He provided side-by-side comparisons of Galileo’s drawings with modern photographs taken at the same lunar phases, and these provide striking support for his contention that Galileo did indeed observe, and record, very accurately.

Whitaker had examined the copperplate engravings of the first edition of Galileo’s *Siderius Nuncius* and seven manuscript images, and he commented on “previously unnoticed differences between the manuscript and printed versions of *Siderius Nuncius*... In the original edition these engravings present a reasonably well-executed appearance, but subsequent editions utilize woodcuts, and the quality deteriorates very rapidly to the point where they are virtually unrecognizable as Moon images. Some of the disparaging remarks made about the drawings undoubtedly stem from examinations of these cruder images”.

M. W. Friedlander

Department of Physics, Washington University, St Louis, Missouri 63130, USA

1. Kemp, M. *Nature* **401**, 116 (1999).

2. Whitaker, E. J. *Hist. Astron.* **9**, 155–169 (1978).

Caught in the act

Forest life encapsulated in amber 20 million years ago can be reassembled.

The Amber Forest: A Reconstruction of a Vanished World

by George Poinar Jr
and Roberta Poinar
Princeton University Press: 1999. 239 pp.
\$29.95, £18.95

S. Blair Hedges

It is hard to imagine a better means of preservation than entombment in amber. Hardened tree resin has preserved organisms ranging from bacteria to mammals for millions of years. Amber surpasses even the finest grain sediments in its ability to retain details, including cellular organelles and pigmentation. Deposits are found throughout the world, but amber from Hispaniola — the island occupied by the Dominican Republic and Haiti — is especially fossil-rich and contains the fossil record of terrestrial life in a tropical environment.

The Amber Forest reconstructs this 20 million-year-old ecosystem, using behaviours of living species to infer habits of their extinct relatives. With more than 200 photographs (most of them in colour) the authors walk the reader through this ancient rainforest, describing a multitude of ecological interactions. In some cases, these interactions were captured vividly in amber: a pseudoscorpion is caught in a face-off with an ant, a resin bug attacking a stingless bee, and a queen ant carrying a scale insect in her mandibles.

The complexity of this lost world is remarkable. Take the spiders, for example. The Dominican amber reveals that these predators had their fair share of specialized enemies, then as now. Spider wasps pull on the strands of a cobweb to entice a spider, then sting it and carry it back to the nest as a food morsel. The larvae of certain beetles impale spiders with a sharp tail spine and release digestive fluids over their bodies. Then there are the spider flies, whose larvae will crawl up an arachnid's leg, bore into the body cavity and consume the living spider from the inside.

The amber came from the resin of one species of tree, the algarrobo, but it preserves a cross-section of the forest. Besides the abundant insects, the amber forest is filled with bacteria, protozoans, fungi, snails, nematodes, tardigrades, onychophorans, crustaceans and many other groups. The plant fossils alone provide evidence for all layers of the forest, from the high canopy species down to the bushes and bryophytes.

If all these could be preserved in amber for millions of years, why couldn't DNA survive the same way? Michael Crichton acknowledged the authors of this book, George and Roberta Poinar, for this idea in *Jurassic Park*. But, though studies have since been published claiming DNA preservation in amber, none has been replicated and researchers remain sceptical. Nonetheless, a negative result should not be interpreted as positive evidence that DNA does not exist in ancient amber. As the authors point out, failure to find it may represent a technological limitation that could be overcome in the future.

Biogeography is a weak aspect of this book. The authors claim that the Hispaniolan biota was mostly carried with the island by plate tectonics from earlier geological connections with North and South America. However, the literature in Caribbean biogeography is extensive and

most evidence favours overwater dispersal as the primary mechanism. In part, this is based on the unusual composition of the modern and fossil (including amber) vertebrate fauna, which agrees more with dispersal than with the wholesale transport of a continental biota.

Finally, it is hard to avoid the topic of conservation in a book about tropical rainforests, even ancient ones. The authors note that tropical rainforests today are disappearing at an alarming pace. This is especially true for the island of Hispaniola, where only a small fraction of the original forest remains and most species probably have yet to be discovered. *The Amber Forest* clearly demonstrates the antiquity of tropical rainforests and gives us yet another reason to mourn their loss.

S. Blair Hedges is in the Department of Biology, Pennsylvania State University, University Park, Pennsylvania 16802, USA.



Amber's contents

If you buy or are given a piece of amber containing an insect or plant, how do you know what you are looking at? Although there are many characteristics that can help to distinguish between the various groups of organisms, most of these are difficult to see in amber because the insect or plant is not in its natural resting state. Help is at hand. Andrew Ross, who is curator of fossil arthropods in the Department of Palaeontology at the Natural History Museum in London, has produced a little book about ways to identify the creatures and plants you are most likely to see captured in amber. *Amber* (Harvard University Press, \$12.95) provides detailed keys to identification,



so that you can easily recognize the lacewing (*Neuroptera*) shown above or the cypress twig (*Thula*), inset.

Knowledge on knowing

The MIT Encyclopedia of the Cognitive Sciences

edited by Robert A. Wilson
and Frank G. Keil

MIT Press: 1999. 964 pp. \$105, £93.50

Manfred Fahle

How do you describe, on a single page, a book such as *The MIT Encyclopedia of the Cognitive Sciences*, with its 471 articles and 132-page introduction? It can be done quite briefly as follows: the main body of the book has 900 pages, the list of contributors fills 11 pages and the indexes occupy some 50 pages. The book was published this year, its size is 29 × 22 × 5.5 cm and it weighs 2.9 kg.

But what are the cognitive sciences? A formal definition of the term is not readily found in the encyclopaedia. Cognitive science is mentioned only once in the index (relating to the essay on Sigmund Freud). However, in the six, long, well-written essays that form most of the introduction, the editors portray the six domains they consider to be most important for the cognitive sciences at the end of the twentieth century: philosophy, psychology, neuroscience, computational intelligence, linguistics and what the editors call "culture, cognition and evolution".

The ambitious aim is to represent, in a single volume, the full range of concepts, methods and findings in cognitive science from the past quarter-century, and to lay the basis for fostering knowledge and understanding between those working in the various fields of cognitive science. The editors certainly achieve this. Furthermore, they want the reader to understand the interconnections between ideas and the links within cognitive science. It is extremely difficult, perhaps impossible, to achieve such integration of information and insight, but the introductory essays and intensive cross-references help. The book combines the advantages of an encyclopaedia, namely succinct, accurate information on a specific term written by specialists, with those of a textbook, conveying more general concepts and a deeper understanding of the structures in a field. The extensive bibliography and cross-references greatly enhance the book's usefulness.

Merely browsing through the tome is enjoyable, and the small bites in which the information is delivered invite reading. While hardly noticing, you learn that blindsight was first described by the Swiss neurologist L. Bard in 1905, that until the eighteenth century most researchers thought the cerebral cortex was simply a protective rind, that in 90 per cent of the adult population language functions are located mainly in the left hemisphere, and much more.

A minor drawback, in my opinion, is that all the entries are of about the same length, so that roughly the same space is devoted to the explanation of, say, psychophysics, illusions or behaviourism as to, for example, Leonard Bloomfield, Paul Grice or Franz Boas. In the latter cases, I would have been happy with just a few sentences, since information is more easily accessible if it is ordered according to topics rather than people.

In conclusion, this is an important encyclopaedia, with a wealth of information supplied by first-rate experts. It also makes agreeable reading. But most people will use it as a database rather than a reading book, as is customary with encyclopaedias. Maybe MIT Press would consider publishing the excellent 100 or so pages of the introduction as a separate, inexpensive booklet.

Manfred Fahle is at the Center for Cognitive Sciences, Human-Neurobiology, Bremen University, Argonnenstrasse 3, D-28211 Bremen, Germany.

A blooming renaissance

The Rose's Kiss: A Natural History of Flowers

by Peter Bernhardt

Island Press: 1999. 267 pp. \$24.95

Peter G. Kevan

Snottygobble may sound like an expletive, but it's not. It's actually a genus of plant (*Persoonia* spp.) in Australia. It has small fruit with gelatinous greenish flesh relished by children who chew it and try to out-disgust each other by slobbering the resultant goo. It is also called the 'honeypot flower' for its chalice-like blooms which have copious nectar. In Peter Bernhardt's eloquent, accurate and literary book, floral biology — or anthecology — is explained to inform and entertain everyone, from curious scientists to amateur natural historians. Botanical science is laced with much zoology. The book describes the fascinating ways in which insects exploit flowers for nectar and may or may not service plants by pollination. The tiny, pollen-pilfering, FuManchu bee (*Leioproctus filamentosa*; Colletidae) has long, slender palpi resembling the villain's moustache and used to rob the snottygobble of its nectar without the bee's affecting pollination. The true pollinators are larger *Leioproctus* spp. and other bees.

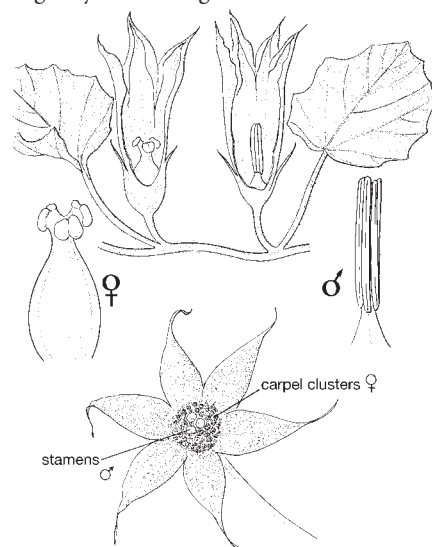
Botany and floral biology are exciting subjects. That excitement is manifest in the rebirth of anthecology in the past 25 years. As it excited Christian Konrad Sprengel 206 years ago to question the functional morphology of flowers, and the Victorian-era naturalists of Europe and America when the

theory of evolution was young, anthecology is again front and centre in many studies in evolutionary ecology, animal and plant reproductive biology, physiology and behaviour, biogeography, biodiversity and conservation.

With provocative cryptic titles and introductory quotations, the seventeen chapters take readers from the basics of floral form and function, plant sex, through the life and death of flowers, and more plant sex. Bernhardt ranges from common temperate-zone woodland herbs, shrubs and trees to rare and strange inhabitants of tropical forests and deserts.

The colour, shape, size and smell of floral attractants are related, with floral rewards such as nectar, pollen, oils and brood sites, to pollinator behaviour and resource requirements. Pollinator diversity is simply yet extensively explored by Bernhardt, from beetles, butterflies, moths, flies and bees to birds and bats noisily squawking in trees. In a chapter called "F is for Fake (and Flower)", the elements of deceitful plants that attract pollinators with misleading signals (mimicry of colour and scent, often coupled with absence of rewards such as nectar) are richly described. The aroids and orchids are notorious at duping pollinators, as are the largest of all flowers, *Rafflesia*, locally known as corpse flowers because of their characteristic smell of rotting meat.

The much understudied, yet ecologically important process of wind pollination is well explained by reference to grasses, trees of temperate countries and hay fever. Next comes more plant sex, but by self-fertilization ('self-made marriages'), and plant celibacy by seeding without fertilization ('virgin births'). The final chapter, by cogently discussing recent discoveries in



F is for flower: male and female squash blossoms grow on the same vine.

palaeontology, addresses Charles Darwin's "abominable mystery" as to the origin of flowering plants (Angiospermae).

This book introduces basic principals to readers who may know little about the natural history of flowers, while presenting scientists, steeped in anthecology, with fascinating insights, twists and a rare glimpse of the personal curiosity that has motivated Bernhardt to write another book.

I congratulate Bernhardt and his publishers on an excellent book. It would make a splendid gift and is a marvellous read. ■

Peter G. Kevan is in the Departments of Environmental Biology and of Botany, University of Guelph, Guelph, Ontario N1G 2W1, Canada.

Research as 'normal'

Die Deutsche Forschungsgemeinschaft in der Weimarer Republik und im Dritten Reich – Wissenschaftspolitik in Republik und Diktatur 1920–1945

by Notker Hammerstein

Beck: 1999. 582 pp. DM98

Marco Finetti

For a long time — far too long — a veil of silence lay over German science and research under the Third Reich. After 1945, many scientists suppressed or hushed up their involvement in the National Socialist regime and instead made careers for themselves in the new German republic. Even subsequent generations showed little interest in examining the past — out of loyalty to their teachers, disinterest, or as a means of protecting themselves. And for many years, the history of science was a subject more or less ignored by (West) German historians.

At last, however, a change seems to be taking place. Young historians, in particular, are beginning to examine the mainly cultural and humanistic research assignments that were carried out under the Nazis. Scientific associations are starting to talk openly for the first time about their activities before 1945 and their refusal to examine these activities since. The Max-Planck-Gesellschaft has commissioned a committee of historians to investigate the history of its predecessor, the Kaiser-Wilhelm-Gesellschaft, under National Socialism. And an initial study has recently been published on the role of the biggest, most influential German scientific organization, the Deutsche Forschungsgemeinschaft (DFG), in the Third Reich.

Because of its unusual past, this study has been awaited with particular anticipation. When the DFG celebrated its seventy-fifth anniversary in 1995, it decided not to mark the occasion with a big jubilee celebration.

Instead, at the request of former president Wolfgang Frühwald, the DFG commissioned an independent historian, Notker Hammerstein, of the Goethe University, Frankfurt, to examine its activities between 1933 and 1945.

The results of Hammerstein's efforts have now been published, and they are disappointing in more than one respect. One reason, although not the most important one, is explained in the foreword by the author himself. Hammerstein declares that he was more interested in other subjects and was not overjoyed at the prospect of taking on the assignment. This becomes plain enough in parts of the book.

But Hammerstein has plenty to say. His history of the DFG becomes a history of German research policy and sponsorship as a whole. And because his study starts, not in 1933 when Hitler seized power, but in 1920 with the foundation of the DFG, he sheds light not only on changes and differences, but also on continuities between the republic and the dictatorship.

In the Weimar Republic, the DFG played fundamentally the same role as it plays today, albeit on a different scale and under different circumstances. As a self-governing body that allocated state grants on the basis of independent reports, it became the most important meeting point for researchers from all disciplines and a driving force behind German research. Its fair and democratic principle was obviously fundamentally opposed to the National Socialists. But the new rulers did not break up the DFG. They used it, as Hammerstein puts it, "to centrally organize university, research and scientific policy in the Third Reich". But the organization came increasingly under the influence of the Reichserziehungsministerium (REM), founded in 1934. In the end, says Hammerstein, the DFG was no more than a "clearing house" for the REM. So, for this reason alone, it had nothing more in common with today's DFG.

The first of two key theories in Hammerstein's book is that the REM exerted a far greater influence on the DFG and Nazi research policy than was previously assumed. The second concerns, not the institutions, but the subjects of research — a far more contentious area.

According to Hammerstein, German research after 1933 and well into the war years was carried out for the most part as it always had been. Researchers in Germany were generally concerned with the same subjects as their colleagues abroad, and most of the DFG-sponsored projects met with the "standards of scientific research that were internationally customary at the time". Naturally, he says, new materials, substances and preparations were also developed and used by the Nazi rulers for their war weapons. But strictly speaking, claims Hammerstein, this research could have been part of "conven-

tional academic endeavours" and might also have helped the sick, for example. Hammerstein does not believe that the scientists, and certainly not the DFG as sponsors, can be blamed for the fact that the results of this research were used for other purposes.

In a similar way, the author discusses another particularly delicate subject — the medical experiments carried out on humans in concentration camps, assignments that were also sponsored by the DFG. Hammerstein does not want to play down these criminal experiments — but he does everything to deny a direct connection between such activities and the DFG. Whether it was research into tuberculosis, twins or hereditary diseases, he claims that the applications were always concerned with the "classical subjects of anthropology and medicine". The fact that humans were used and killed in such experiments, he says, was not made clear in the research applications. Hammerstein almost seems to suggest that the misuse of the DFG and its funds for criminal purposes has a dual use — albeit of a horrifying kind.

Hammerstein's theory about the normality of German research after 1933 has already triggered heated debates in Germany. And it deserves particularly critical examination not least because it is the result of a certain methodical procedure. Allowing sources to speak for themselves may be an accepted principle of historical science. In Hammerstein's case, however, this principle takes the place of interpretation and personal opinion in parts of the study, or even replaces them entirely. Thus, we get the impression that a professional historian has fallen victim to his sources. One of many examples: Hammerstein repeatedly describes how many German researchers deceived themselves into believing that they were able to carry out 'normal' research even under the Nazis, thereby succumbing to such self-deception himself.

This impression is endorsed by several linguistic oddities and oversights. For example, Hammerstein writes of the "bad" experiments, or of the "notorious" Josef Mengele, which sounds rather clumsy or naive in a scientific paper. Elsewhere, he mentions Jewish mathematicians who were brought together from various concentration camps to carry out important war research — according to Hammerstein, this was a "comparatively harmless use of concentration camp prisoners". One may ask again whether it is really appropriate to compare everything.

This is an ambivalent book, which raises more questions than it answers. Here, too, we can agree with Frühwald, the initiator of Hammerstein's study, who claims the book has sparked off but has by no means ended the debate about the DFG's role in the Third Reich. ■

Marco Finetti was editor-in-chief of the *Deutsche Universitätszeitung* in Bonn and is now a scientific writer in Cologne.

Science in culture

The Story of Time

An exhibition at the National Maritime Museum, Greenwich, and an accompanying book (£25)
J. L. Heilbron

The National Maritime Museum's millennial exhibition opens this week at the Queen's House in Greenwich, where the world's day begins. The display of hundreds of costly books, pictures, instruments and other artefacts was underwritten by *The Times*, Parmigiani Fleurier (makers of fine timepieces) and, primarily, J. P. Morgan, the round-the-clock bankers.

Time is more than money. The exhibition offers five major themes — the creation, measurement, depiction, experience and end of time — with examples from everywhere and everywhen. The section on creation includes drawings from *Genesis*, a tenth-century BC Babylonian clay tablet, an Aztec stone god from about 1500 AD, a magnificent Navajo sand painting dated 1966, bronze and clay statuettes of the Hindu god Shiva from the seventeenth and eighteenth centuries, and a nineteenth-century Maori wooden door lintel.

The section on time measurement enlarges the coverage to Islam (represented by eleventh- to thirteenth-century brass astrolabes), Japan (ivory zodiacal animals, nineteenth century), ancient Rome (Mithraic marble personifying time, second to third centuries) and China (Father Time (pictured) in jade and porcelain, seventeenth and nineteenth centuries).

The Story of Time, the lavishly illustrated book of the exhibition, describes 400 items — two-thirds Western European and the rest from a variety of other cultures. The book reflects the structure of the exhibition: it is a visually appealing, culturally varied, imaginative, multimedia hotchpotch. The director of the Royal Greenwich Observatory, Kristen Lippincott, who conceived and curated the exhibition, wants it to deliver the message that "the watch is not the truth". She means not only that the division of time into hours and minutes is arbitrary, but also that clock-time embodies concepts and experiences peculiar to the modern, European-dominated world. Lippincott hopes to open new vistas — and perhaps also new truths — to her visitors. She is most persuasive. Visitors intent on catching a particular train home, however, will find it convenient to retain their belief in watch truth.

Interspersed in the catalogue are 23 essays on time in art, chronometry and history. On the arts side are



Time, Chinese-style: the wooden strips above form one half of a calendar for the year 63 BC, while the mirror to the left is decorated with cosmological symbols.

essays about portraits, nineteenth- and twentieth-century art, music and *vanitas* images, filled with skulls and other reminders of life's transience. On the technical side are essays on calendrical principles, horology and the concept of time in different cultures — among geologists, for example, and Inuits. All these accounts are valuable, though not all make good use of the images around them and

some may be too compressed to be fully accessible to general readers.

Three essays call for special mention, two because their authors, Umberto Eco and E. H. Gombrich, appear on the title page, and the third, by Felipe Fernández-Armesto, because it is a particularly elegant

response to a difficult assignment. Eco's purpose, in his essay called "Times", is to point out "some areas of confusion" faced or created by analysts of time. What was God doing before Day One? Is the future before or behind us ('the weeks ahead', 'the weeks to follow')?

Do we imagine the past spatially to our left or to our right? How long will the world last? Eco might have added: "How did Joshua

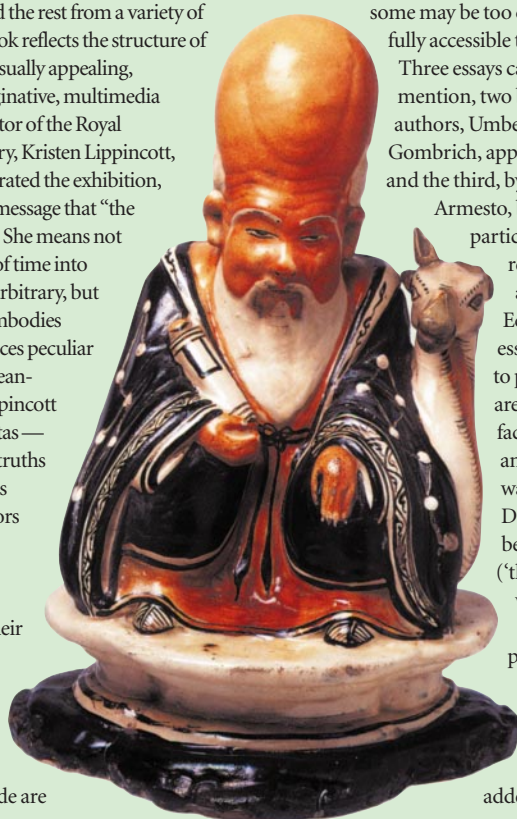
tell the time while the Sun stood still?"

Most people will be more concerned with the pressing question of the longevity of the Universe. *The Story of Time* gives three different answers computed from the same Hindu recipe. Perhaps none is right. This much is sure, however: the world will be full of pitfalls no matter how long it lasts. Eco falls into one himself. He writes, "For millennia, the only reliable clock was the cock's crow." Is this not Euro- or even Mediterranean-centrism? As John MacDonald observes in his excellent essay "Inuit time", the Inuits — who live much of their lives in the dark and do not keep chickens — use the orientation of the circumpolar stars, the wake-up cries of their infants and the pressure in their bladders to reckon the start of the day to within rooster accuracy.

Gombrich originally composed "The history of anniversaries" in 1974, for the centenary of the philosopher Ernst Cassirer's birth. He calls attention to the historical ingredients in the concept of centennial: a widely disseminated and long-lived calendar, a corporate memory, and agreement on things worth celebrating. Recognizing that years ending in two zeros are worth noticing required the division of history by century and a consensus that 100 years is a good period for stock-taking. The first century turn that met these criteria was 1800.

Fernández-Armesto's assignment was to discuss "Time and history". His answer: "time is history's subject matter and history [is] time's diet". This cannibalism has had three forms: a cyclical one, in which nothing singular can happen, since everything will happen again, and no one bothers with dates; a linear one, in which the story has a beginning, an end and dates in between; and a chaotic one, characteristic of our time, in which historians have embraced "the counter-factual, the self-reflexive, the representational, the random, the causeless, the unverifiable, the liminal and the implicit". If so, the end of history has come before the end of time.

Although the exhibition might best be described as fitting the chaotic view of history, it is well worth the visit. Its venue is a monument of architecture. Its artefacts are handsome and well presented. Its catalogue is attractive and informative. Together, they indicate the excellence achievable at the end of our century by up-to-date museology inspired by good ideas, enriched by good scholarship and enabled by just plain money. J. L. Heilbron is at Worcester College, University of Oxford, Oxford OX1 2HB, UK.



Annus physicalis 1932

A bumper crop of physical discoveries — something in the (heavy) water?

H. B. G. Casimir

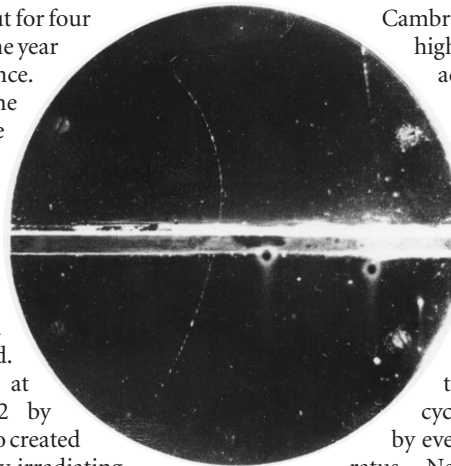
A physicist reviewing our century will find a discovery or new insight in every year, but 1932 stands out as a true *annus mirabilis*. John Dryden, who coined the term to describe 1666, discussed three events: the plague, the fire of London, and the war with the Dutch. The physicists' harvest for 1932 yielded four major discoveries: the neutron, heavy hydrogen, nuclear reactions induced by accelerated particles, and the positive electron, or positron.

Was there anything mysterious and special about 1932? I do not think so, but during the 1920s study of the behaviour of electrons in atoms led to the creation of quantum mechanics. The time was ripe for physicists to go further — it is not surprising that a number of fundamental discoveries were

made in the 1930s, but for four of them to come in one year is just a matter of chance.

The existence of the neutron, a particle without electric charge and with roughly the same mass as the nucleus of the hydrogen atom (the proton), had been postulated by Ernest Rutherford. It was discovered at Cambridge in 1932 by James Chadwick, who created beams of neutrons by irradiating beryllium with alpha particles. It was at once clear that the existence of the neutron led to a simple model for nuclear structure. Nuclei were henceforth regarded as compounds of protons and neutrons. The simplest case is one proton and one neutron. This particle was called deuteron, D. Its oxide, D₂O, is the molecule of heavy water. It was found in 1932 that about one part in six thousand of normal water is heavy water.

Because neutrons have no charge, they are not repelled by the electric fields of other nuclei. Therefore they can enter them and produce all manner of nuclear reactions. By giving a proton sufficient energy — in effect, using brute force — it is also possible for it to overcome electrical repulsion. This was done by John Cockcroft and Ernest Walton, also at



Cambridge. With a 700-kilovolt high-voltage generator they accelerated protons and let them fall on lithium.

They showed that some of the lithium nuclei broke in two — and proudly announced that they had split the atom. So began the era of big machines. The high-voltage generators were succeeded by cyclotrons, the cyclotrons by even more powerful apparatus.

Neutron, deuteron and accelerators formed the beginning of an extremely active period of nuclear physics that led to nuclear energy and weapons.

The fourth discovery, the positron, was made by Carl Anderson, using what could be called the astronomical method — examining phenomena outside our control, in this case the so-called cosmic radiation.

Here we see clearly the difference between physics and astronomy. The physicist can arrange and manipulate the things he or she studies, the astronomer cannot. Astronomers can observe in many ways, using all the resources of physics and technology, and may even build an observatory on the Moon, but they cannot influence the structure or behaviour of planets, let alone stars. Nor can they influence the cosmic radiation.

Apart from the sign of its charge, a positron is identical to a normal electron. Entirely new, however, is that a positron-electron pair can be created by collision processes in empty space, and that the pair can disappear. The partners annihilate one another, leaving their energy behind as electromagnetic radiation. Again, the positron's existence and behaviour were predicted, this time by Paul Dirac.

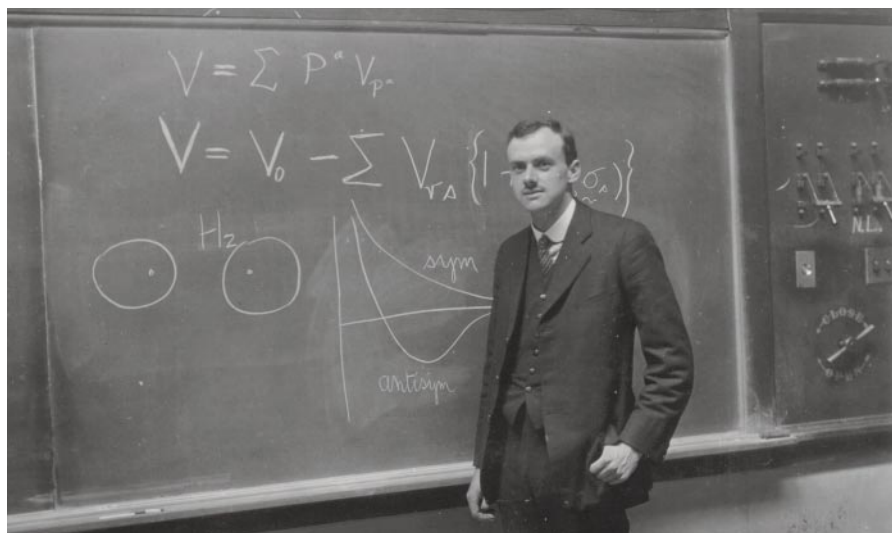
The positron was the first short-lived particle to be created out of empty space. It is a prelude to the high-energy physics that blossomed after the Second World War. Today, ever more complicated machines grant us fleeting glimpses of whole families of short-lived particles, many of them predicted by theory. Sometimes it almost appears that the theories are not a description of a nearly inaccessible reality, but that so-called reality is a result of the theory.

In any case, the limits of this strange new world are not yet in sight.

H. B. G. Casimir is at De Zegge 7, 5591 TT Heeze, The Netherlands.



Positive images: the first positron track (top), discovered by Anderson (left), confirmed Dirac's predictions (below).



COMP.BASILISK FAQ

Frequently asked questions about basilisks.

David Langford

1. What is the purpose of this newsgroup?

To provide a forum for discussion of basilisk (BLIT) images. Newsnet readers who prefer low traffic should read comp.basilisk.moderated, which carries only high-priority warnings and identifications of new forms.

2. Can I post binary files here?

If you are capable of asking this question you MUST immediately read news.announce.newusers, where regular postings warn that binary and especially image files may emphatically not be posted to any newsgroup. Many countries impose a mandatory death penalty for such action.

3. Where does the acronym BLIT come from?

The late unlamented Dr Vernon Berryman's system of maths-to-visual algorithms is known as the Berryman Logical Imaging Technique. This reflected the original paper's title: "On Thinkable Forms, with Notes towards a Logical Imaging Technique" (V. Berryman & K. Turner, *Nature* **409**, 340–342; 2001). Inevitably, the paper has since been suppressed and classified to a high level.

4. Is it true that science-fiction authors predicted basilisks?

Yes and no. The idea of unthinkable information that cracks the mind has a long SF pedigree, but no one got it quite right. William Gibson's *Neuromancer* (1984), the novel that popularized cyberspace, is often cited for its concept of 'black ice' software which strikes back at the minds of hackers — but this assumes direct neural connection to the net. Basilisks are far more deadly because they require no physical contact.

Much earlier, Fred Hoyle's *The Black Cloud* (1957) suggested that a download of knowledge provided by a would-be-helpful alien (who has superhuman mental capacity) could overload and burn out human minds.

A remarkable near-miss features in *The Shapes of Sleep* (1962) by J. B. Priestley, which imagines archetypal shapes that compulsively evoke particular emotions, intended for use in advertising.

Piers Anthony's *Macroscopic* (1969) described the 'Destroyer sequence', a purposeful sequence of images used to safeguard the privacy of galactic communications by erasing the minds of eavesdroppers.

The comp.basilisk community does not want ever again to see another posting about

the hoary coincidence that *Macroscopic* appeared in the same year and month as the first episode of the British TV programme *Monty Python's Flying Circus*, with its famous sketch about the World's Funniest Joke that causes all hearers to laugh themselves to death.

5. How does a basilisk operate?

The short answer is: we mustn't say. Detailed information is classified beyond Top Secret.

The longer answer is based on a popular-science article by Berryman (*New Scientist*, 2001), which outlines his thinking. He imagined the human mind as a formal, deterministic computational system — a system that, as predicted by a variant of Gödel's Theorem in mathematics, can be crashed by thoughts that the mind is physically or logically incapable of thinking. The Logical Imaging Technique presents such a thought in purely visual form as a basilisk image which our optic nerves can't help but accept. The result is disastrous, like a software stealth-virus smuggled into the brain.

6. Why 'basilisk'?

It's the name of a mythical creature: a reptile whose mere gaze can turn people to stone. According to ancient myth, a basilisk can be safely viewed in a mirror. This is not generally true of the modern version — although some highly asymmetric basilisks like B-756 are lethal only in unreflected or reflected form, depending on the dominant hemisphere of the victim's brain.

7. Is it just an urban legend that the first basilisk destroyed its creator?

Almost everything about the incident at the Cambridge IV supercomputer facility where Berryman conducted his last experiments has been suppressed and classified as highly undesirable knowledge. It's generally believed that Berryman and most of the facility staff died.

Subsequently, copies of basilisk B-1 leaked out. This image is famously known as the Parrot for its shape when blurred enough to allow safe viewing. B-1 remains the favourite choice of urban terrorists who use aerosols and stencils to spray basilisk images on walls by night.

But others were at work on Berryman's speculations. B-2 was soon generated at the Lawrence Livermore Laboratory and, disastrously, B-3 at MIT.

8. Are there basilisks in the Mandelbrot set fractal?

Yes. There are two known families, at symmetrical positions, visible under extreme



OLIVER BURSTON

magnification. No, we're not telling you where.

9. How can I get permission to display images on my website?

This is a news.announce.newusers question, but keeps cropping up here. In brief: you can't, without a rarely granted government licence. Using anything other than plain ASCII text on websites or in e-mail is a guaranteed way of terminating your net account. We're all nostalgic about the old, colourful web, and about television, but today's risks are simply too great.

10. Is it true that Microsoft uses basilisk booby-traps to protect Windows 2005 from disassembly and pirating?

We could not possibly comment.
(Revised 27 June 2006)

David Langford, a former weapons physicist, has won many of science fiction's Hugo awards for criticism and commentary in the field.

The book of genes

Peter Little

The sequence of human chromosome 22, now published, is the first phase in a biological revolution. Not least, the result will be a transformed appreciation of human individuality.

As 1999 draws to a close and we approach the third millennium, a new book is being written. It will change the way we see ourselves as profoundly as did the momentous books of the first two millennia — the great books of religion and the *Origin of Species*. The first chapter of this book, the book of genes, is summarized on page 489 of this issue¹. It consists of the DNA sequence of most of human chromosome 22 and contains the inaugural results of the Human Genome Project. The whole of the book will probably never be printed in its entirety — it would take about half-a-million pages of this journal to do so — but it will nevertheless inexorably alter our perceptions of human health, attitudes to each other and understanding of our uniqueness. It is fitting that in this age our most telling monument, a culmination of some of the great biological discoveries of our time (Fig. 1), may be an electronic database.

The DNA in all normal human cells is in 23 pairs of pieces, neatly packaged into 46 chromosomes, and the paper summarizes the understanding we have of the 33.4 million or so base pairs that make up the DNA sequence of just one of these pairs (number 22, chosen because it is one of the smallest human chromosomes; only chromosome 21 is smaller). The summary is brief. There are certainly 545 genes encoded in the DNA and there may be as many as 1,000 (see refs 2,3). The sequence is not quite complete because for technical reasons there are 11 gaps; however, and some are just a few thousand. The approach taken, and an alternative one, are described in Box 1 (overleaf).

The sequence tells us about biology in several ways. We do not know whose DNA it is, so we do not learn about a specific person. Instead we draw lessons that apply to any human. All of the constituent parts of the sequence have been generally available for many months^{2,3} — indeed, as fast as they came off the analysers, they were put into free databases — and so the information has already been used by at least seven groups to study newly discovered genes.

There are thought to be at least 27 human disorders associated with changes to genes on chromosome 22, and the causative genes in eight of them have still to be discovered. The conditions range from cancers to disorders of fetal development and of the nervous

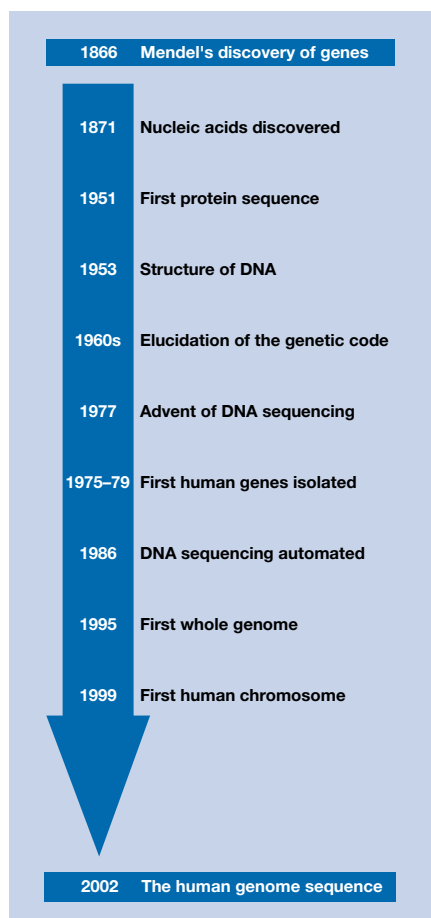


Figure 1 Sequence of events — from Mendel's discovery of the existence of genes, to 2002 and the projected date for completion of sequencing the human genome. The first protein to be sequenced was insulin; the first genome, that of the bacterium *Haemophilus influenzae*.

system. The biological jewel in this particular crown is probably the nature of a gene involved in schizophrenia, thought to be located within the sequence but, as yet, unidentified. The sequence also provides tantalizing clues as to how chromosome 22 must have evolved because at least eight regions of the DNA are present in duplicate — a characteristic of evolution of the genetic apparatus. Frustratingly, however, we cannot read its history without comparing its organization with that of the DNA of close animal relatives of humans, and we do not yet have this information.

So what is so revolutionary about all of this? In itself, perhaps not much. But what

book can be appreciated by reading only one chapter? Rather, it is as a foretaste of what is to come that the paper¹ is a milestone in our understanding of our own biology. Within three years, the full sequence of human DNA will be finished and for the first time we will be able to identify the proteins, some 200,000 to 300,000 of them, that are used to make a human being.

The idea that a profound biological problem can have bounds is not part of our usual thinking; all such problems are so complex that research normally touches upon only part of them at any one time. Soon, however, we will have the bounds created by the complete list of proteins that might underpin the process we study. Of course, being bounded is not necessarily the end of the story — the idea that the Atlantic is bounded is no real comfort to anyone who might want to swim across it — and the difficulties of using this protein list effectively are very great indeed. Until now biologists have mainly relied upon intuitive models of how biological systems develop and respond. The new protein world will move the models into areas of analytical complexity where few researchers are trained to go.

A good example of such complexity is in the systems that signal from the outside of a cell to the inside⁴, the end result often being gene activation or deactivation. Individually, one can think of these systems as light circuits with three components — a switch, a wire and a light. The three components have just two interactions, switch to wire and wire to light. If two such signalling circuits can interact (switch with switch, wire with wire, and light with light) there will be 11 interactions. For three circuits, there will be 27 interactions, and so on. The increase in interactions is dramatically nonlinear, and in the face of that intuition is valueless. How will we meet the challenge of the potential interplay of complete sets of proteins drawn from a list that has perhaps a thousand entries?

In more practical terms, what is going to happen next? Where will the DNA sequence take us? First, it will take us to the list of proteins, but not quickly. DNA encodes proteins via RNA, and we have known this code for 30 years. But this does not mean a gene is easy to identify in DNA. Of the 1,000 or so possible genes on chromosome 22, only 545 are easy to spot because they encode proteins that are

similar to ones that have been studied previously in organisms as disparate as humans and bacteria. The remaining, putative genes are predicted by complex computer modelling that is only partly accurate. Showing that these are, or are not, genes will be a continuing challenge.

Second, will we know the functions of these proteins just because we know their sequence? In most cases we will not, but biologists will have huge numbers of proteins and families of proteins to think about and experiment on. At least by a process of exclusion, we will know what we do not know because the list will, in the end, be complete.

This property opens up the third path that the DNA sequence will take us down — the prospect of global analysis of gene activity — that is, the amounts of RNA each gene is, or is not, producing. This is possible using 'arrays' and 'chips' for comprehensive assay of the variation in RNA levels, a wholly novel approach to biology⁵. We do not know what it will tell us about human beings. But the results⁶ of studying just 15,000 genes are already striking, and provide a completely new viewpoint of the life of a cell.

Finally, the DNA sequence will become the framework upon which we will be able to place the changes to the sequence which are present in every one of us. Any two unrelated human beings differ by one base pair in every 1,000 or so. A base pair is chemically defined by the nucleotides it contains and so these differences (polymorphisms) are called single-nucleotide polymorphisms or SNPs (often called 'snips'). About 85% of all DNA differences are SNPs and the population biology of the human species has meant that they are very common — hence the one-in-1,000 difference⁷.

There is a general consensus that SNPs are probably the cause of most common genetic disorders. We all carry many SNPs but if we are unlucky enough to carry the 'wrong' set of changes, we are predisposed to one or other of the common disorders with a genetic component such as diabetes, heart diseases, asthma or cancers⁷. The 'SNP consortium', a group funded by The Wellcome Trust and key pharmaceutical companies, is setting out to find 300,000 of the commonest SNPs in human populations. Once this has been done, we can confidently face the task of identifying which ones contribute to the common disorders — or indeed to any common human trait, be it good or bad.

The reason that each of us is an individual is because we are the products of genes, gene variance and life experiences ('environment'), a combination that defines our unique personalities, abilities and disabilities. If knowledge of gene differences can be combined with an understanding of the richness of environmental influences, we will have the key to unlocking the cause of most of the common disorders that kill or

Box 1 : Two routes to the genome

There is controversy over the best way to sequence the human genome. The public sequencing project, of which the consortium that sequenced chromosome 22 is a member, is jointly supported by the US National Institutes of Health and Department of Energy, and the UK Wellcome Trust. It uses a 'clone-by-clone' approach. In contrast, Celera Genomics, a private company in the United States, has set out to sequence the whole human genome directly by the 'random-shotgun' method.

Clone-by-clone sequencing is sequencing applied to 150,000-base-pair fragments of human DNA (the clones), one at a time. The entire sequence of a chromosome is then reassembled. The random-shotgun approach simply goes directly to the totality of our DNA by sequencing at random. Why the two strategies? The human genome is full of duplications of DNA sequences and contains about a million 'interspersed repeats'. These are short, near-identical copies of the same sequence, and they make assembly of DNA sequences difficult because the wrong repeats can get matched together when sequences are being assembled by computer. Advocates of the clone-by-clone technique argue that

working on one DNA clone at a time simplifies the repeats problem by concentrating on just those that are found in a single clone. Backers of the random-shotgun approach believe that the problem can be overcome with massive computing power.

At the time of writing, the fruitfly genome — about one-tenth the size of the human — has nearly been completed by Celera using random sequencing. This is sometimes held to be a model for the human project, but it is not: not only is the fruitfly genome smaller, it is much less rich in interspersed repeats and duplications. There will probably be very many gaps in the final human sequence assembled by the random approach.

The arguments for and against^{8,9} the two techniques are not easy to reconcile. One advantage of the clone-by-clone approach is that useful information is generated well before completion of the entire sequence, as has been shown with chromosome 22 (ref. 1). But the work that goes into the clone maps is considerable and taxing (and there are still ten small gaps in the chromosome 22 map). In March of this year the public consortium announced that it would

sequence 90% of human DNA by spring 2000. The goal will be achieved by continuing a clone-by-clone approach, as before, but not completing each 150,000-base-pair fragment. The result will be a 'draft' sequence that will not be continuous, in that each 150,000 base pairs will be contained in 10,000–15,000 base-pair pieces. This is less useful than the fully 'finished' sequence achieved for chromosome 22 but will still be enormously powerful for writing the book of genes. The project will then move into a finishing phase, and has a projected completion date of 2002.

Who is right? We do not know but can hope that both sets of data can be combined to give us the very clearest picture of the genome. One crucial difference remains: only the public consortium is making the DNA sequence freely available, as it comes off the machines, to anyone who wants it. The consortium's 'Bermuda agreement' is a binding set of rules for immediate and free release of sequence, and is to be applauded — it would be a terrible blow for science and humanity if the human genome became a commercial property. **P. L.**

otherwise cause suffering.

This is the field of genetic epidemiology, the analysis in large populations of the interactions of gene, genetic variation and environment. Its study will require new techniques for DNA testing, new statistical tools for analysing hundreds of thousands of people and the utmost care to protect individual privacy. The goal is worth attaining because genetic epidemiology will contribute to discovering new drug treatments, will enable those treatments to be tailored to individual genetic and life profiles, and will predict some of our futures and contribute to the advice needed to avoid the unpleasant ones. It will change the emphasis of health care from treatment to prevention. But ethical dilemmas will also arise as we discover some of the sources of the differences in individual health, behaviour and personality.

Biology is entering a new world; not only

do we face a revolutionary leap in what we know, we also face radical changes in the tools we must use to understand that information. I am not sure that we are prepared for the full impact of either but we have already made our first tentative steps into the new world of the genome. The challenge is now to translate the new biology into tangible benefits for humanity. ■

Peter Little is in the Department of Biochemistry, Imperial College, Prince Consort Road, London SW7 2AZ, UK.
e-mail: p.little@ic.ac.uk

1. Dunham, I. *et al.* *Nature* **402**, 489–495 (1999).
2. <http://www.sanger.ac.uk/HGP/Chr22>
3. <http://www.genome.ou.edu/human.html>
4. Weng, G. *et al.* *Science* **284**, 92–96 (1999).
5. "The Chipping Forecast" *Nature Genet.* **21** (Suppl.) (1999).
6. Iyer, V. R. *et al.* *Science* **283**, 83–87 (1999).
7. Chakravarti, A. *Nature Genet.* **21** (Suppl.) 56–60 (1999).
8. Green, P. *Genome Res.* **7**, 410–417 (1997).
9. Weber, J. L. & Myers, E. W. *Genome Res.* **7**, 401–409 (1997).

Mid-ocean ridges

One more time, from the top

Erik Hauri

One of the commonest geological processes on Earth is the partial melting of rocks. It is responsible for the creation of Earth's crust, and largely occurs at the network of undersea volcanoes that make up the mid-ocean ridge system. From studies of these ridges and their fragments now exposed on land (ophiolites), a well-accepted model for the creation of oceanic crust has emerged. As tectonic plates spread apart at mid-ocean ridges, the underlying mantle rocks (called peridotites) rise up to fill the gap, and decompression during this upward flow causes the mantle to melt and erupt the basalt magmas that form the overlying oceanic crust.

On page 514 of this issue¹, however, Benoit *et al.* report evidence of a very different type of magma (andesite) from an ophiolite in Oman. They suggest that andesite could be generated by repetitive melting of seawater-altered rocks at the very top of the mantle at a mid-ocean ridge, and implicate a process quite different from normal melting.

The results of ocean-based studies provide a good explanation for the chemical composition of mid-ocean ridge basalt (MORB): decompression melting of mantle peridotite, followed by crystal separation during cooling of MORB magma. The result is a layered crustal structure, depicted in all three parts of Fig. 1, in which the MORB lavas overlie a coarse-grained layer of crystals (gabbroic cumulates) with a layer of magma feeder channels (dikes) in between.

Geochemical analyses have further shown that the mantle rocks at the top of the mantle, just beneath the oceanic crust, have undergone large amounts of melting, removing as much as 30–35% of the original mass as MORB magma at high temperatures^{2,3}. As a result, the top of the mantle at mid-ocean ridges is very depleted in elements that preferentially partition into basalt magma, such as sodium (Na), titanium (Ti), strontium (Sr), and light rare-earth elements (LREE) such as neodymium (Nd). In consequence peridotite rocks at the top of the mantle can generate little additional magma, even if reheated to high temperatures; and the magma that does form would have very low concentrations of Na, Ti, Sr and LREE, and have a very different overall chemical composition to typical MORB. The extremely depleted chemical signature would also be transferred to any crystals that formed from such a magma.

It is precisely this chemical signature that Benoit *et al.*¹ have found in a subset of gabbroic rocks, called D (for depleted) gab-

bronorites, from the Oman ophiolite. They show that the depleted gabbros also contain unusually large amounts of the mineral orthopyroxene, which they suggest can only be the result of crystallization from a magma (andesite) with much more silica than normal MORB basalts.

How can such depleted magmas be generated beneath mid-ocean ridges? Presumably, the top of the mantle is so depleted in Na, Ti, Sr and LREE that it cannot yield additional magma. Nonetheless other examples of such rocks exist — as the authors say, gabbroic rocks with similarly depleted signatures have been recovered in drill cores from a site (DSDP 334) on the mid-Atlantic ridge, and low-Na plagioclase crystals similar to those in the Oman D-gabbro-norites are a very minor but widespread component of otherwise normal MORB samples. Studies of the site 334 gabbros⁴ and low-Na plagioclase crystals⁵ have suggested that they are the crystal products of magmas formed at the very end of the melting process that pro-

duces normal MORB. The final fraction of magma formed at the top of the mantle during decompression melting would be depleted in Na, Ti, Sr and LREE because of the preceding extraction of magma earlier in the melting event.

Rather than accept this model, Benoit *et al.*¹ have appealed to a second melting event to create the Oman D-gabbro-norites. The authors went to great effort to separate clean, unaltered pyroxene crystals from these rocks and determine their ⁸⁷Sr/⁸⁶Sr isotopic ratio (this ratio is a direct tracer of the composition of the magma source because, unlike element abundances, it is unchanged by the melting process). The results are surprising — the Sr isotope ratios of the N (normal) gabbro-norites are similar to those in typical MORB magmas, but those of the D-gabbro-norites are much higher. This measurement indicates that the highly depleted rocks were not formed by separation of magma fractions late in the melting process. Rather, they must have been formed in a separate event involving a different peridotite source to that which generates normal MORB.

To identify this source, the authors look to another fundamental process — the circulation of sea water through the oceanic crust at regions of hydrothermal activity. As it seeps deeper into the cooling oceanic crust, sea water becomes heated to the point where it will react with the surrounding rocks, resulting in their partial dissolution and the precipitation of hydrothermal minerals. In the process, Sr is exchanged between sea water and the crustal rocks, and the very high Sr isotope ratio of sea water can be imparted to those parts of the crust that become highly altered. It is thought (but not known) that sea water may seep all the way through the seven kilometres of crust to alter even the top of the mantle.

Benoit *et al.* propose that hydrothermal circulation adds water and seawater-derived Sr to the highly depleted peridotite at the top of the mantle. The water decreases the melting temperature of the peridotite, which as a result melts easily during conductive heating resulting from later decompression melting events. That is, the mantle initially melts by decompression melting, yielding typical MORB, and the top of the mantle becomes highly depleted in Na, Ti, Sr and LREE. The top of the mantle then undergoes hydrothermal alteration by sea water, and is re-melted by subsequent conductive heating. This feature of the model stems from the restriction of the D-gabbro-norites in the Oman ophiolite to the margins of a large mantle intrusion — an observation that could have been made only at an ophiolite, where the geological features of the oceanic plate are clearly exposed.

One nagging uncertainty about this model remains. The highly depleted chemical signatures of the D-gabbro-norites have

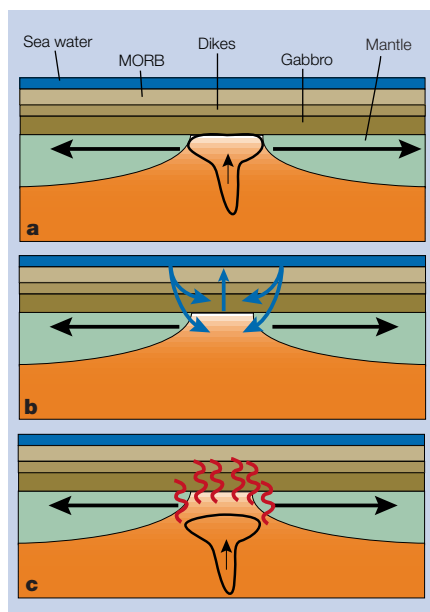


Figure 1 General structure of ocean crust (a–c), and the mantle-alteration model of Benoit *et al.*¹. a, Decompression melting produces a layered oceanic crust, and an underlying mantle that is highly depleted in sodium, titanium, strontium and light rare-earth elements at the top. b, Hydrothermal circulation of sea water (blue) adds water and seawater strontium to the top of the mantle. c, Later mantle intrusions heat the altered mantle, forming highly depleted, silica-rich melts which crystallize in veins within the oceanic plate (red). MORB, mid-ocean ridge basalt.

never been recognized in actual lava samples from the mid-ocean ridges, and only rarely as isolated inclusions of melt trapped within crystals⁶. Perhaps this process is not common and the highly depleted magmas are mixed with large amounts of normal MORB during eruption. If so, then how important is this process in the larger view of magmatism at mid-ocean ridges?

The answer may lie in further study of the site 334 gabbros and the widespread low-Na plagioclase crystals that are sometimes found in normal MORB. In particular, melt inclusions trapped in low-Na plagioclase may have retained other signatures of sea water hydrothermal alteration, such as increased water and chlorine content. Oxygen isotope studies of these and similar samples could also confirm the role of water in

the generation of highly depleted ocean-ridge magmas. Meantime we are left to ponder the involvement of hydrothermal circulation in repetitive melting of the mantle, and its influence on magma composition at mid-ocean ridges. ■

Erik Hauri is in the Department of Terrestrial Magnetism, Carnegie Institution of Washington, 5241 Broad Branch Road NW, Washington DC 20015, USA.
e-mail: hauri@dtm.ciw.edu

1. Benoit, M., Ceuleneer, G. & Polvé, M. *Nature* **402**, 514–518 (1999).
2. Dick, H. J. B., Fisher, R. L. & Bryan, W. B. *Earth Planet. Sci. Lett.* **69**, 88–106 (1984).
3. Johnson, K. T. M., Dick, H. J. B. & Shimizu, N. *J. Geophys. Res.* **95**, 2661–2678 (1990).
4. Ross, K. & Elthon, D. *Nature* **365**, 826–829 (1993).
5. Meyer, P. S. & Shibata, T. *ODP Sci. Res.* **106/109**, 123–142 (1990).
6. Sobolev, A. V. & Shimizu, N. *Nature* **363**, 151–154 (1993).

Alzheimer's disease

A firm base for drug development

Bart De Strooper and Gerhard König

Alzheimer's disease is a progressive dementia in which massive deposits of aggregated protein-breakdown products — amyloid plaques and neurofibrillary tangles — accumulate in the brain. The amyloid plaques are thought to be responsible for the devastating mental decline in patients with Alzheimer's disease, and the main component of these plaques is a small peptide called amyloid β ($A\beta$). This peptide is made when its precursor, the amyloid precursor protein (APP), is cleaved, and two distinct proteases are thought to do this job. The β -secretase cleaves at the β -site (Fig. 1), whereas the γ -secretase cuts APP at the γ -site¹.

The identity of these secretases has long been sought, but half of the mystery, at least, is now over. Papers in *Science*² and *Molecular and Cellular Neuroscience*³, as well as reports by Yan *et al.*⁴ and Sinha *et al.*⁵ (on pages 533 and 537 of this issue), describe the identification of the β -secretase. And these exciting results could, perhaps, accelerate the development of drugs to stop or even reverse the neurodegenerative process.

Development of any medication requires, among other things, validated animal models and well-defined molecular targets. For research into Alzheimer's disease we lacked both just a few years ago. Then disease-causing mutations were discovered in the APP and presenilin genes, allowing animal models to be developed⁶. But none of the disease genes were the type of classical drug target that drug hunters enjoy chasing⁷. Proteases such as the secretases, by contrast, would be excellent pharmacological targets — take HIV research, for example, where the identification of viral proteases in the mid-1980s

led to the successful development of marketed protease inhibitors about ten years later⁸.

Four groups (not surprisingly, all from the pharmaceutical and biotechnology industries) have now identified a convincing candidate ' β -site APP-cleaving enzyme' (BACE), also known as Asp-2. Vassar and colleagues² inserted, in a screening-like fashion

ion, pools of unknown human genes (complementary DNAs) into a cell line that they knew expressed their target protease. Of 860,000 clones, Vassar *et al.* identified just one candidate protease (BACE) that increased the release of $A\beta$ from their chosen cells. Naturally occurring APP mutations at the β -site cause early-onset Alzheimer's disease, and the authors found that mutant APP was a better substrate for BACE than normal APP. In an impressive series of cell-biological experiments, Vassar and colleagues went on to show that BACE fulfils all possible criteria for a candidate β -secretase.

Yan and colleagues⁴ took a very different approach. Based on pharmacological data, these authors knew that the β -secretase must have aspartyl-protease-like characteristics — acidic pH requirements and a signature sequence (aspartic acid-serine/threonine-glycine) found in most aspartyl enzymes. Using this information, Yan *et al.* first scanned the genome of the nematode worm *Caenorhabditis elegans*, which has been almost completely sequenced, for candidate proteases. They then went back to the vertebrate genome and identified homologues of the ten hits they obtained. Several candidates were further validated by extensive cellular (antisense) experiments. This approach (which is often known to produce false-negative results) allowed

them to identify Asp2 as the β -secretase, using the same criteria as Vassar and colleagues.

It is very reassuring that BACE has also been identified by a classical protein-fractionation

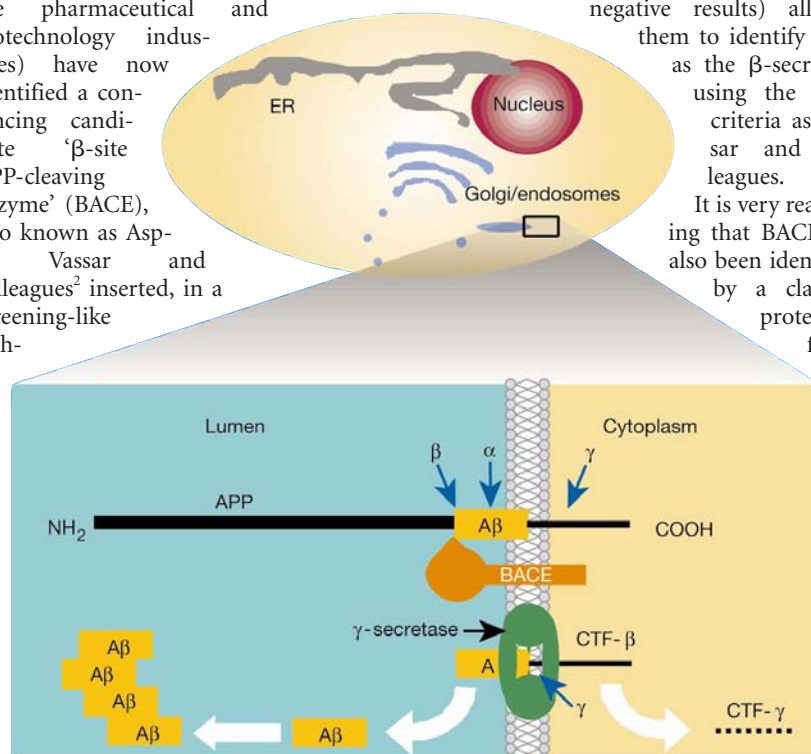


Figure 1 Proposed mechanism by which the amyloid β ($A\beta$) peptide is generated through cleavage by the newly identified β -site APP-cleaving enzyme, BACE^{2–5}. The APP molecule is inserted into the lumen of the endoplasmic reticulum (ER) and then transported to the cell surface. Either in the Golgi, or in early endosomes after endocytosis, APP is cut by the transmembrane-spanning BACE. This cleavage results in secreted soluble APP (β -sAPP) and the membrane-associated carboxy-terminal fragment β (CTF- β). The CTF- β is a substrate for the γ -secretase (model adapted according to ref. 9), which cleaves its substrate in a symmetrical fashion in the transmembrane domain. This results in secreted $A\beta$ and rapidly degraded CTF- β .

tionation approach. Sinha and colleagues⁵ established that the β -secretase is a membrane-associated aspartyl protease with an optimal pH of 5.5. They then observed, like the other groups^{2,4}, that this presumably aspartyl-like activity is unusual in that it is not inhibited by pepstatin, a typical aspartyl-protease inhibitor. But to purify BACE, Sinha and colleagues needed a molecular hook. They designed several variants of the APP sequence, spanning the β -site, including so-called transition-state analogues. Such analogues 'freeze' a bound protease in the act of cleaving the substrate. Sinha *et al.* then tested their analogues against crude preparations of human brain containing abundant BACE activity. They found that one amino-acid substitution, from aspartate to valine with a statine analogue, at position +1 (that is, on the carboxy-terminal side of the cleavage site) resulted in a potent inhibitor with a half-maximum inhibition at 30 nM. Using this molecular hook, Sinha and colleagues pulled out their candidate protease from human brain extracts, with a 300,000-fold enrichment.

Finally Hussain and colleagues³, like Yan *et al.*, named their β -secretase Asp2 — suggesting they have more than one candidate protease. However, Hussain *et al.* did not reveal in their paper how they obtained their cDNA clone from their proprietary expressed sequence tag (EST) database. They did show, however, that point mutations in Asp2 (or BACE) at both of its two active sites (the aspartic acid-serine/threonine-glycine sequence) mean that it can no longer process APP to A β .

Ever since it became clear that proteases chop APP down to A β , these proteases have been prime targets for drug discovery. In the absence of molecular targets, cellular reporter systems have been used to develop compounds that reduce the amount of A β produced, and we may soon see the first clinical trials of these drugs. But the isolation of BACE means we can now screen for drugs that act directly on the target protease. Future structural information from X-ray diffraction studies of BACE with a bound inhibitor might give valuable insights into the design of new structural classes of inhibitors.

Several challenges remain, however. The BACE and its homologue BACE-2 belong to a new class of membrane-bound aspartyl proteases. Are there other BACE homologues? And, if so, will these have to be considered for selectivity screens in drug-optimization studies? We also do not know which other precursor molecules or cellular processes depend on proper BACE activity. Transgenic mice with these genes knocked out, either conditionally or totally, will be very useful for resolving such questions. Another problem is the subcellular location of BACE, in the lumen of the Golgi body and endosomes (Fig. 1). This means that

inhibitors will have to cross at least two lipid bilayers — a formidable penetration hurdle for even small-molecular-weight compounds. Moreover, any BACE inhibitor has to pass the blood-brain barrier to find its target in neurons. New compounds will therefore need to have excellent pharmacokinetic properties.

Despite all of this, the identification of the β -secretase means that the path towards specific inhibitors is now set, and it is time not only to test the amyloid hypothesis ('*in vivo veritas*'), but to find a way of halting this dreadful disease.

Bart De Strooper is at the Flanders Interuniversity

Institute for Biotechnology and K. U. Leuven, B-3000, Leuven, Belgium.

e-mail: Bart.Destrooper@med.kuleuven.ac.be

Gerhard König is at Bayer AG, Pharma Research CNS, 42096 Wuppertal, Germany.

e-mail: gerhard.koenig.gk@bayer-ag.com

1. Haass, C. & Selkoe, D. J. *Cell* **75**, 1039–1042 (1993).
2. Vassar, R. *et al.* *Science* **286**, 735–741 (1999).
3. Hussain, I. *et al.* *Mol. Cell. Neurosci.* www.academicpress.com/www/journal/cn/mcne
4. Yan, R. *et al.* *Nature* **402**, 533–537 (1999).
5. Sinha, S. *et al.* *Nature* **402**, 537–540 (1999).
6. Price, D. L. *et al.* *Annu. Rev. Genet.* **32**, 461–493 (1998).
7. Drews, J. *Nature Biotechnol.* **14**, 1516–1518 (1996).
8. Ren, S. & Lien, E. J. *Prog. Drug Res.* **51**, 3–31 (1998).
9. Lichtenthaler, S. F. *et al.* *Proc. Natl Acad. Sci. USA* **96**, 3053–3058 (1999).

Chemical physics

Ultrafast relaxation in water

Abraham Nitzan

When a molecule is excited, where and how fast does its excess energy go?

The answer to this question is a prerequisite for understanding and predicting the course of many chemical and biological processes. Chemical reactions can take place only when enough internal energy has accumulated in the molecule. The study of chemical reactions is therefore intimately connected with the study of energy-relaxation processes that compete with the chemical reaction channel. Intermolecular excitation transfer is one such process that has been studied for more than half a century. A paper by Woutersen and Bakker on page 507 of this issue¹ may force us to re-examine some of

our notions about this important relaxation pathway.

When an isolated atom is optically excited it can relax to the ground state only by emitting radiation. In large molecular systems many degrees of freedom compete for the excitation energy and the winner rarely takes all. After a molecule is excited in solution, much of its energy usually ends up as increased solvent thermal motion, and can be regarded as wasted. Chemical interest often lies in other relaxation channels, for example electron transfer or chemical bond breaking. But even pathways that eventually lead to wasted thermal energy, such as intermolecular excitation transfer, can be of great interest at intermediate timescales. Knowing how energy flows between different molecular modes *en route* to complete relaxation can suggest ways to direct it to more useful channels, similar to harnessing the water flow in a river for useful work.

Intermolecular excitation energy transfer is the process by which one excited molecule, a donor, transfers its excess energy to another, an acceptor, leaving the latter in an excited state. This process continues until terminated by photon emission, chemical reaction or thermal relaxation. The intermolecular excitation pathway can be desirable or not, depending on your objective: it may obstruct an attempt to bring a molecule into higher excited states and it will destroy coherence that may otherwise help control a photochemical reaction. On the other hand it can provide ways to 'conduct' energy to where it is needed; an example is the use of sensitizers in photographic films in order to activate photoreactions in species that do not absorb natural light. Nature has also learned how to use these processes, for example in the light-harvesting complexes of photosynthetic systems².

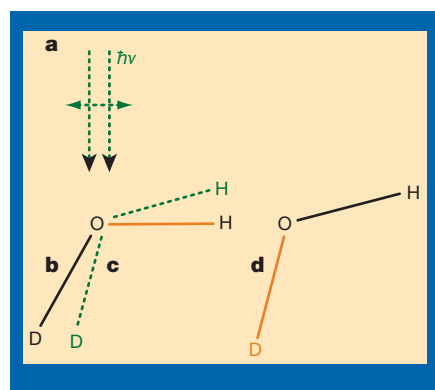


Figure 1 Intermolecular energy transfer. Woutersen and Bakker¹ measure energy transfer in pure H₂O and in mixtures of HDO dissolved in D₂O using two infrared pulses. A pump pulse, polarized as shown in a, excites the OH bond of the HDO molecule, b. The induced rotational anisotropy can relax either by rotation of this molecule to configuration c, or by energy transfer to another molecule, d. Averaging over all final orientations amounts to loss of anisotropy, which is monitored by the probe pulse (not shown).

The theory of such energy-transfer processes goes back to the well-known works of Förster³ and Dexter⁴. The simplest Förster transfer mechanism is similar to the interaction between two electric dipoles. The rate of energy transfer, k , is described by $k = T^{-1}(r_0/r)^6$, where T is the lifetime of the excited state, r is the distance between the donor and acceptor, and r_0 is a parameter called the Förster radius. This equation tells us that the rate of dipolar energy transfer behaves like r^{-6} . With increasing r , higher-order interactions (such as dipole–quadrupole, quadrupole–quadrupole and exchange interactions) decay much more rapidly than the dipole–dipole interaction, and are effective only at very small intermolecular distances.

How important is this mode of energy flow? It is significant only when its rate is comparable to or faster than other relaxation processes. The most important competing processes are intramolecular vibrational relaxation, where vibrationally excited molecules relax by transfer of energy within the molecule itself, and vibrational energy relaxation, where vibrational energy is transformed into solvent thermal energy. These processes are fast; relaxation of polyatomic molecules in condensed phases at room temperature occurs over a few picoseconds or less.

In contrast, vibrational energy transfer between molecules is generally believed to be too slow to be important. It is only expected to play a significant role for diatomic molecules or at cryogenic temperatures, where vibrational relaxation is relatively slow. Indeed, these are the conditions under which such processes have been observed in the past^{5,6}. Contrary to such expectations, the experiment by Woutersen and Bakker¹ shows that resonant intermolecular energy transfer between OH bonds in liquid water is extremely fast. Moreover, it appears to be much faster than the vibrational energy relaxation of the OH group, which has recently been shown to have a short lifetime of 740 fs (femtoseconds)⁷.

In their experiment, Woutersen and Bakker use two 200-fs infrared pulses: one relatively strong, linearly polarized ‘pump’ pulse to excite the OH groups and another, low-intensity pulse to probe this excitation. They use thin-layer samples of either pure water (liquid H₂O), or a mixture of HDO and D₂O (D is deuterium, a heavy isotope of hydrogen). The mixed samples make it possible to measure the dependence of the energy transfer rate on the OH concentration. Woutersen and Bakker measure the rotational anisotropy of the molecules as a function of the time delay between pump and probe. Rotational anisotropy is induced by the pump pulse, which excites molecules with specific orientations. This vibrational excitation can then relax either by rotational

motion of the excited molecules or by energy transfer between molecules of different orientations (Fig. 1).

These two relaxation modes can be distinguished from each other for the low OH mixtures (that is, HDO dissolved in D₂O): energy transfer between HDO molecules depends on the concentration of this species, whereas their rotation does not. Therefore measuring the rotational anisotropy at different delay times as a function of HDO concentration yields the characteristic time for molecular rotation (four picoseconds), and more importantly the rate of intermolecular vibrational energy transfer between the OH groups. These results show that the Förster theory accounts well for the observed intermolecular vibrational energy transfer in HDO–D₂O mixtures and that the corresponding transfer rate is quite fast — in the range of a few picoseconds for molar concentrations of OH. With Woutersen and Bakker’s technique, the transfer rate is measurable even though the competing processes of energy relaxation are very fast.

The real surprise comes from similar measurements in pure H₂O. Using the Förster results from mixtures of HDO in D₂O ($r_0 = 2.1$ Å) to extrapolate to the intermolecular distance in pure water (2.8 Å) predicts an energy-transfer time in the range of a few hundred femtoseconds. But the observed intermolecular energy transfer in pure water takes place even faster than the experimental time resolution of ~ 100 fs. This is considerably faster than the 740-fs lifetime of the excited OH population, and makes intermolecular vibrational energy transfer one of the fastest relaxation processes ever recorded in water. This means that vibrational energy cannot be localized in water long enough to affect most chemical reactions. On the other hand it implies that water is an extremely good conductor of vibrational energy through its OH groups. It is even possible that this energy-transfer process could involve other molecules containing OH groups, so water may play an important role in protein dynamics when energy is transported between different molecules.

The failure of the dipolar Förster theory in H₂O is not unexpected because the OH groups are so close to each other that higher-order interactions come into play. But the observation that intermolecular vibrational energy transfer in water is so amazingly fast calls for a reassessment of vibrational energy transfer and relaxation in condensed phases. Many questions remain. For example, what is the actual rate of vibrational energy transfer in water? Is this behaviour peculiar to this liquid (perhaps it is associated with its special structural properties)? Previous studies from the same laboratory⁷ have suggested that the fast vibrational energy relaxation of the OH bond in water is possibly associated



100 YEARS AGO

Two or three months ago reports were published in the daily press of the discovery of an electrical method of giving sight to the blind. It was alleged that Mr. Stiens had succeeded in constructing an electrical apparatus which performed all the functions of the eye and was an efficient substitute for it. Like many other newspaper reports of so-called scientific discoveries, this has proved to be without sound foundation. Mr. G. H. Robertson, who is himself afflicted with blindness, describes in the *Electrician* the results of personal inquiries into the matter with a member of the staff of our contemporary. In spite of several visits to Mr. Stiens, no experimental proof in substantiation of the claims which were put forward on his behalf was obtained, and the conclusion arrived at is that these claims are groundless. Life is so short and crowded with so many important duties that it is impossible to investigate the many sensational statements made by irresponsible interviewers, but we are grateful to any one who will take the trouble to examine some of the rumours which are put forward in the name of science.

From *Nature* 30 November 1899.

50 YEARS AGO

In British Astronomical Association Circular No. 312 some details are given regarding the two newly discovered satellites of Uranus and Neptune, respectively. Both were discovered by Gerard P. Kuiper during his search for new satellites with the 82-in. reflector of the McDonald Observatory, University of Texas. The new satellite of Uranus, now named Miranda, was discovered on February 16, 1948, magnitude 17, and is now known to have a period of about 33h. 56m. The motion is approximately circular and in the plane of the other four satellites. Neptune ii, for which the name Nereid has been proposed by the discoverer, was found on May 1, 1949, on plates exposed for forty minutes at the prime focus, with the mirror stopped down to sixty-six inches ($f/5$). Its magnitude was estimated to be 19.5, and later observations show that its period is about two years and that the plane of its orbit is within six degrees of the ecliptic. Kuiper says that, as Neptune could retain satellites nearly ten times as far away as Nereid, with periods up to about fifty years, further work is planned to cover the outer regions of the system.

From *Nature* 3 December 1949.

with its coupling to the nearest hydrogen bond. Is there a link between that and the present observations? If not, what is the origin of this extremely fast energy-transfer process? The need to answer these questions is our next challenge. ■

Abraham Nitzan is at the School of Chemistry, Sackler Faculty of Science, Tel Aviv University, Tel Aviv 69978, Israel.
e-mail: nitzan@post.tau.ac.il

Neurobiology

Derailed axons get on track

Kai Zinn and Aloisia Schmid

How do growing axons in the central nervous system navigate through the dense jungle of cells and processes that they encounter on the way to their targets? On page 540 of this issue, Bonkowsky *et al.*¹ show that the choice of pathways for growth cones (the leading edges of growing axons) in the fruit fly *Drosophila melanogaster* is regulated by a receptor tyrosine kinase known as Derailed (Drl). These results further indicate

1. Woutersen, S. & Bakker, H. J. *Nature* **402**, 507–509 (1999).
2. van Oijen, A. M., Ketelaars, M., Köhler, J., Aartsma, T. J. & Schmidt, J. *Science* **285**, 400–402 (1999).
3. Förster, Th. in *Modern Quantum Chemistry* Vol. III (ed. Sinanoglu, O.) 93–137 (Academic, New York, 1965).
4. Dexter, D. L. *J. Chem. Phys.* **21**, 836–850 (1953).
5. Legay, F. in *Chemical and Biochemical Applications of Lasers* Vol. II 43–86 (Academic, New York, 1977).
6. Apkarian, V. A., Wiedeman, L., Janiesch, W. & Weitz, E. *J. Chem. Phys.* **85**, 5593–5610 (1985).
7. Woutersen, S., Emmerichs, U., Nienhuys, H.-K. & Bakker, H. J. *Phys. Rev. Lett.* **81**, 1106–1109 (1998).

that, rather than attracting growth cones to the right pathways, Drl causes them to be repelled from the wrong ones.

The array of axons in the embryonic *Drosophila* central nervous system has a ladder-like structure (Fig. 1). Anterior and posterior commissural tracts cross the midline in each body segment, and two longitudinal tracts extend the length of the embryo. Each of the roughly 300 neurons within a unit of this structure is thought to extend its axon along a genetically determined pathway. Most interneurons (neurons that synapse with other neurons) extend axons across the midline of the central nervous system, and attractive and repulsive factors that regulate this fundamental crossing decision have been identified^{2,3}. But axon guidance at the midline involves more than just the decision

whether or not to cross. Each commissure contains many distinct pathways, and the growth cone of a particular neuron always follows the same one. The unique sequence of navigational decisions made by its growth cone determines the complex and invariant shape of each axon in the central nervous system⁴.

Although we are far from understanding how complete axon trajectories (such as those in Fig. 1) are determined, the results of Bonkowsky *et al.*¹ are an important step forward. These authors define how axons choose between the two major subdivisions of the crossing pathways — the anterior and posterior commissures. They show that Drl is normally expressed on axons that follow the anterior commissure, but not on those that take the posterior route^{1,5}. Moreover, forced expression of Drl on specific axons that normally take the posterior commissure causes them to choose the anterior tract instead (Fig. 2).

Can Drl redirect any crossing axon into the anterior commissure? To address this question, Bonkowsky *et al.* simultaneously misexpressed the Commissureless (Comm) protein and Drl on a set of axons that normally never cross the midline (the thoracic Ap axons). Expression of Comm in neurons downregulates a repulsive signal from the midline that is transduced by the Roundabout (Robo) receptor. So, non-crossing axons that express Comm are diverted into pathways that cross the midline⁶. The authors found that when the thoracic Ap axons expressed only Comm, they crossed in the posterior commissure. But when they

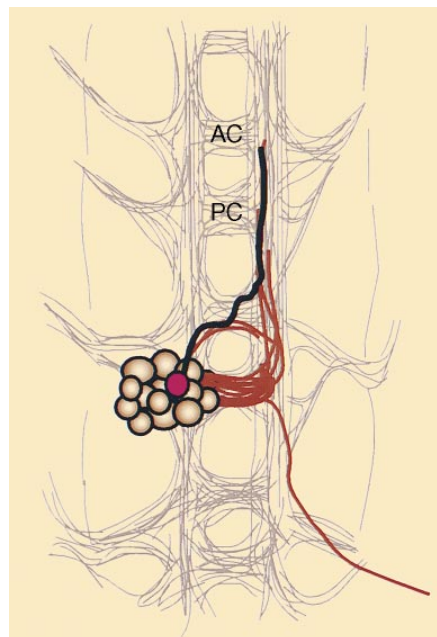


Figure 1 Individual axons follow complex trajectories within the central nervous system (CNS). The axon scaffold of the fly CNS is drawn in grey. A specific neural lineage (from neuroblast 5–2) is indicated in brown; the lineage includes neurons that project in both the anterior commissure (AC) and the posterior commissure (PC). One AC neuron (red) — presumably a cell that expresses Drl — extends its axon (black) along the posterior edge of the AC. At the midline the axon veers anteriorly, grows along a bundle at the anterior edge of the AC, then turns into the longitudinal tract. (Modified from ref. 4.)

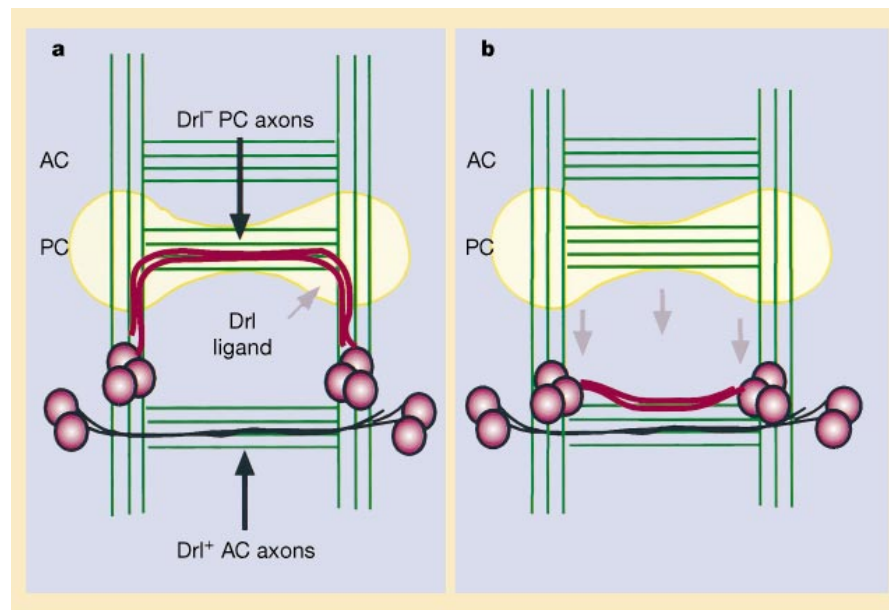


Figure 2 Effects of Drl expression on axon guidance. Bonkowsky *et al.*¹ have found that expression of a protein called Drl controls pathfinding by causing growth cones to be repelled from the wrong pathway. a, Wild-type embryos. The AC (black) and PC (red) Eg axons cross the midline. The distribution of the putative Drl ligand is indicated by yellow shading, and the axon scaffold is in green. b, Forced expression of Drl in the PC Eg axons¹. These axons are repelled by Drl ligand (grey arrows) and instead project across the AC.

expressed Drl as well, they switched to an anterior pathway.

The authors next asked whether Drl regulates commissure choice by making the anterior commissure more attractive, or whether it does so by generating aversion to the posterior commissure. Their data indicate that it repels axons from the posterior commissure. First, when Bonkowski *et al.* expressed Drl without Comm on non-crossing Ap axons that grow anteriorly within the longitudinal tract, the axons were not diverted into the anterior commissure. Instead, their growth cones stalled when they reached the edge of the posterior commissure, suggesting that this region had become repulsive to them.

Second, because Drl is a receptor tyrosine kinase, one would expect it to be activated through an interaction with an extracellular ligand. Although such a ligand has not been identified, Bonkowski *et al.*¹ nonetheless examined its probable distribution by incubating embryos with 'tagged' versions of the Drl extracellular domain, then visualizing these tags using immunohistochemistry or immunofluorescence. The tagged Drl fusion proteins bound to a dumbbell-shaped region of the central nervous system that includes the posterior commissure and adjacent domains within the longitudinal tracts (Fig. 2). These results are consistent with the idea that, during normal development, Drl signalling is activated when growth cones from axons destined for the anterior commissure contact the Drl ligand in the vicinity of the posterior commissure. This repels them from the posterior commissure, allowing them to choose anterior pathways instead.

The next steps will be to identify the ligand(s) for Drl and the signalling proteins downstream of Drl in the growth cone. A more global issue raised by these studies is how growth cones integrate the many attractive and repulsive influences that they meet while navigating through the embryonic landscape. For example, in Fig. 1 consider the growth cone of the black Drl-expressing axon in the anterior commissure. It is attracted to the midline by Netrins (and other factors), but is also influenced by repulsion factors. This repulsion is neutralized by Comm — to allow the growth cone to cross the midline — but is later re-activated to prevent the axon from returning to the midline after it has crossed^{2,3}.

Although Drl repels this axon from the posterior commissure, other factors must attract its growth cone to the anterior pathway. (Drl is not essential for an axon to choose the anterior commissure, and most axons still follow their normal pathways¹ in loss-of-function *Drl* mutants.) The tracts in the fly's central nervous system are composed of discrete bundles of axons that are pioneered by growth cones of early neurons. Each later axon then follows a pathway along

specific bundles. So, the individual pathways across the midline probably correspond to axon bundles or groups of bundles within the commissures. Finally, then, the growth cone must choose among the axon bundles in the anterior commissure and switch its bundle preference at the midline (Fig. 1). Does a growth cone need to integrate all of these attractive and repulsive signals simultaneously? Or are the signals temporally segregated, so a growth cone responds to only one influence at any time?

The Drl repulsion mechanism may be used at other times and places during fly development. For example, mutations in the *Drl* (aka *linotte*) gene affect learning, and adult brain structures implicated in learning cross the midline abnormally in *linotte* flies⁷. The Drl protein is also expressed on body-wall muscles, and growing muscle fibres sometimes overshoot their normal attachment points in *Drl* mutants⁸. Both of these effects imply that motile cells cannot recognize 'stop' signals when Drl is not expressed.

Finally, because Drl is homologous to the mammalian RYK tyrosine kinase⁵, and because other components of midline-crossing pathways are conserved between vertebrates and flies^{2,3}, Drl mechanisms will also probably be involved in development of the vertebrate nervous system. ■

Kai Zinn and Aloisia Schmid are in the Division of Biology, California Institute of Technology, Pasadena, California 91125, USA.
e-mails: zinnk@its.caltech.edu
aschmid@its.caltech.edu

1. Bonkowski, J. L., Yoshikawa, S., O'Keefe, D. D. & Thomas, J. B. *Nature* **402**, 540–544 (1999).
2. Zinn, K. & Sun, Q. *Cell* **97**, 1–4 (1999).
3. Van Vactor, D. & Flanagan, J. G. *Neuron* **22**, 649–652 (1999).
4. Schmid, A., Chiba, A. & Doe, C. Q. *Development* **126**, 4653–4689 (1999).
5. Callahan, C. A., Muralidhar, M. G., Lundgren, S. E., Scully, A. L. & Thomas, J. B. *Nature* **376**, 171–174 (1995).
6. Kidd, T., Russell, C., Goodman, C. S. & Tear, G. *Neuron* **20**, 25–33 (1998).
7. Simon, A. F., Boquet, I., Snguelakakis, M. & Preat, T. *Mech. Dev.* **76**, 45–55 (1998).
8. Callahan, C. A., Bonkowski, J. L., Scully, A. L. & Thomas, J. B. *Development* **122**, 2761–2767 (1996).

Hearing

Spreading the fluid word

Jonathan Ashmore and Jessica de Boer

In *Fantastic Voyage*, the 1966 sci-fi film classic, Racquel Welch and her intrepid team of brain surgeons are shrunk to nanometre dimensions and sent on a journey through the human body. Escaping killer lymphocytes, they get entangled within the cochlea of the inner ear and are buffeted by the sounds as they enter the fluids of the duct. What the leading lady, although appropriately clad in a wetsuit, failed to do was to take

any measurements of the fluid flow in the cochlea. Students of cochlear mechanics have regretted the lost opportunity ever since. Over 30 years later, reporting on page 526 of this issue, Elizabeth Olson¹ has managed to make just such biophysical measurements, with implications for the way we think about models of hearing.

The organs of hearing, on either side of the human head, are coiled compartments

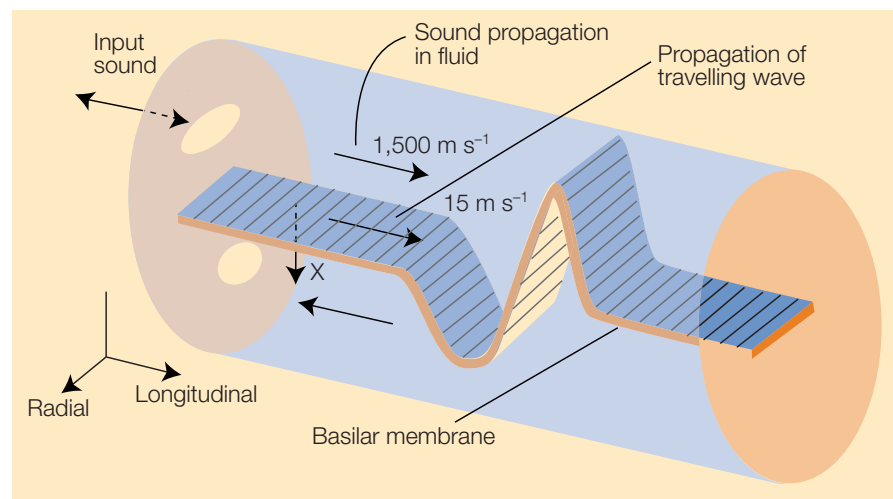


Figure 1 The cochlea uncoiled. The basilar membrane supports a travelling wave. The membrane divides the cochlear duct into two fluid-filled compartments, where sound propagates roughly 100 times faster. Only a central, longitudinal section is portrayed. The vertical excursion of the wave is exaggerated about a million times. The fluid velocity vector at X, perpendicular to the membrane, is the measurement reported by Olson¹.

within the temporal bones. If the coils were straightened out, each cochlea would be 34 mm long in humans and about one-third of that size in small mammals². The basilar membrane—a collagen-fibre membrane no more than 0.5 mm wide—divides the cochlea along its length. This membrane, together with the sensory hair cells of the organ of Corti that also span the length of the cochlea, transmits information about incoming sounds into a pattern code sent to the brain via the auditory nerve. The idea is that each component sound frequency is mapped by the mechanics onto a unique pattern of displacements of the membrane.

Sound waves travelling in air are transmitted to the fluid in the cochlear duct with little loss of energy. But how is the pattern of hair-cell excitation then set up along the basilar membrane? Modelling the vibration pattern of the basilar membrane has been a problem for nearly 140 years—ever since Helmholtz³ first tackled it but ignored the hydrodynamics. The guiding principle since then has been to treat the cochlea as a one-dimensional transmission line⁴. The hydrodynamics implied in a three-dimensional cochlea have been difficult to add in, both computationally and experimentally. Olson's work¹ indicates that three-dimensional fluid flow may have to be included in any complete and realistic description of hearing.

Sound waves propagating in cochlear fluid travel, as in water, at around $1,550 \text{ m s}^{-1}$. Drop a stone into a pond and the sound it creates propagates to the bottom at this speed, although the ripples travel across the surface much more slowly (a Rayleigh wave) at a velocity that depends on the depth of the pond and the surface tension of the water. The basilar membrane is a two-dimensional surface with variable stiffness along its length that separates the two fluid compartments (Fig. 1). It effectively propagates a 'surface' wave—known as the Bekesy travelling wave—with a velocity of about 15 m s^{-1} (ref. 5). This wave starts at the base of the cochlea, nearest the middle ear, and propagates with decreasing wavelength but increasing amplitude towards a position of maximum amplitude beyond which it rapidly decays. In the cochlea of a living animal, the amplitude of the travelling wave at this peak is further increased about 100 times by the action of the outer hair cells that lie along the length of the cochlea^{6,7}.

Olson¹ offers a rare experimental view of the dynamic pressure changes in the cochlea. It is technically very difficult to make mechanical measurements in the cochlea owing to the extremely small size and delicate nature of the structures, especially at the basal end where high frequencies are processed. This has slowed progress considerably, and there are few direct measurements of pressure waves. Olson has over-

come these problems by developing a small but sensitive hydrophone that can be placed into the gerbil cochlea and measure pressure changes within the cochlear fluids.

The new data show that fluid flow perpendicular to the basilar membrane falls away within $15 \text{ }\mu\text{m}$ of the membrane. This result seems paradoxical. It has been supposed (in the so-called 'long wave' limit⁴) that fluid should be carried with the basilar-membrane wave throughout the height of the cochlear duct at the basal end of the travelling wave. Because water is incompressible, conservation of mass may imply that fluid is moving in a radial rather than perpendicular direction to the long axis of the cochlea. Alternatively, this is an indication that previous models of the cochlea may have ignored subtle features of the mechanics at the basal end of the cochlea.

Olson's result is consistent with the idea that the cochlea is a three-dimensional, dynamic structure. Not only does it have length, but the cochlear duct also has width and a variable height that determines the mass of fluid moving with the basilar membrane⁵. There are also indications that the organ of Corti on the membrane must be modelled as a complex matrix that includes not only hair cells, but also other cell types with varied mechanical properties. Modern computing power allows us to show that the structure may alter the fine detail of the local vibration pattern in the basilar membrane⁸. This pattern can involve higher radial vibrational modes, implying that surrounding fluid would be both pushed and pulled across, as well as along, the membrane^{9,10}. It is a huge challenge to design a small enough pressure sensor, compared with the membrane width, that could resolve the dynamic detail required for the cochlea (Olson's is $160 \text{ }\mu\text{m}$ in diameter).

We have yet to construct physiologically based models of the cochlea that convincingly describe all of its real-time abilities to analyse complex sounds. Such models should now probably include local interactions between the compartment fluids and the basilar membrane. It is hard to avoid thinking that cochlear models will get more complex before they become simpler. ■

Jonathan Ashmore and Jessica de Boer are in the Department of Physiology, University College London, Gower Street, London WC1E 6BT, UK. e-mail: j.ashmore@ucl.ac.uk

1. Olson, E. *Nature* **402**, 526–529 (1999).
2. Fay, R. R. & Popper, A. N. *Comparative Hearing: Mammals* (Springer, New York, 1994).
3. Helmholtz, H. *Die Lehre von den Tonempfindungen* (Vieweg, Brunswick, 1863).
4. de Boer, E. in *The Cochlea* (ed. Dallos, P.) 258–317 (Springer, New York, 1996).
5. Lighthill, J. *J. Fluid Mech.* **106**, 149–213 (1981).
6. Ruggero, M. A. et al. *J. Acoust. Soc. Am.* **101**, 2151–2163 (1997).
7. Nobili, R. et al. *Trends Neurosci.* **21**, 159–167 (1998).
8. Kolston, P. *Proc. Natl Acad. Sci. USA* **96**, 3676–3681 (1999).
9. Russell, I. J. & Nilsen, K. E. *Nature Neurosci.* **2**, 848–853 (1999).
10. Cooper, N. P. *J. Physiol.* **520**, S114 (1999).

Daedalus

Happy landings

Aircraft are notorious for not failing safe. If anything goes wrong, it can be a terrible task to keep the plane in the air until it can land at an airport. Ideally, aircraft should be fitted with parachutes for safe descent; but a chute big enough to support a plane would occupy almost its whole interior. Daedalus now has a way out.

A parachutist hauls on the webs of his canopy just before hitting the ground. As a result, the parachute descends a little faster, and the man beneath it a little slower, at the crucial moment of impact. DREADCO engineers are now taking this strategy to extremes. They are devising a small but very strong parachute, to be deployed from an aircraft in distress, from a door in the top of the fuselage. Its horizontal flight will be checked below stalling speed, and it will plummet towards the earth, suspended in a horizontal attitude beneath its chute. The chute, however, is attached to the plane not by a simple anchorage, but by a winch. As the plane falls, the winch unwinds rapidly, paying out the parachute on many hundreds of metres of cable. But just as the plane is about to hit the ground, the winch is powered in reverse, and winds the cable forcefully in again. This hauls the parachute downwards far faster than its terminal velocity. So the parachute exerts a powerful upward force on the plane, which slows at the last minute and hits the ground quite gently.

A reliable safety system must be as simple as possible. As the parachute unreels, Daedalus wants it to wind up a spring on the winch, whose tension will wind it in again when the craft is about to hit the ground. This moment will be determined by a weighted cable lowered from the craft, which will go slack when its far end touches the ground. This purely mechanical system will have no vulnerable electronics or powered units.

At last the worry of flying will be relaxed. In case of decompression, engine failure, hi-jacking or structural fatigue, a plane will be able to drop safely out of the sky onto the friendly land. By angling the plane on its tether, the pilot could even 'glide' it towards the best landing site, as human parachutists do. Even if the plane hit the sea, the gentle impact would leave it undamaged. It would float until rescue arrived.

David Jones

The further Inventions of Daedalus (Oxford University Press), 148 past Daedalus columns expanded and illustrated, is now on sale. Special *Nature* offer: m.curtis@nature.com

Obituary

Thomas Hughes Jukes (1906–99)

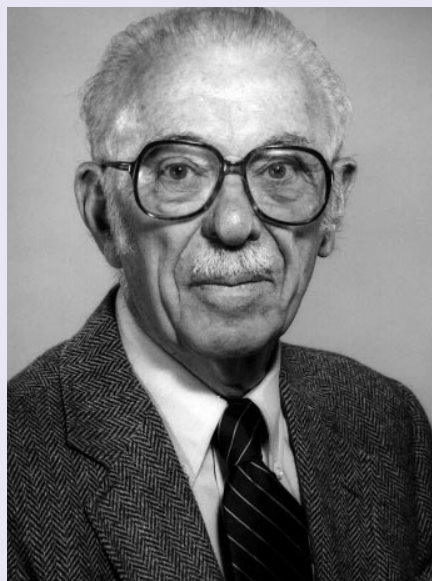
Vigour, both intellectual and physical, was the hallmark of Tom Jukes, who died on 1 November in California. He was habitually coy about his age, now revealed as 93. When asked his age, he would insist that he had been born this century; then, to demonstrate his youth, he would lead his visitors on a 'stroll' over the hills above Berkeley at such a pace that all but he would be out of the breath required to interrupt his running commentary on the state of science and the follies of some of its practitioners. But even before he had an artificial knee fitted in the early 1980s, Jukes did not tempt his visitors to the high Sierras, which were his great love.

He crammed several careers into his long life. Born in Hastings on the English south coast, he took off for Toronto in 1924, where in 1933 he earned a PhD in the department at which Frederick Banting and Charles Best had discovered insulin. His first career was as a nutritionist at the University of California, mostly at the Davis campus where, through feeding experiments with chickens, he was able to unravel the relationships between the different components of the B-vitamin complex. He was one of the authors of the first report that vitamin B₁₂ will cure the deficiency disease pellagra in human beings.

That experience was the platform for Jukes's second career, as an industrial scientist at the Lederle laboratories in New Jersey, initially as head of a programme on nutrition and physiology. There, he discovered the role of folic acid as a vitamin (the lack of which is now believed to be linked with the emergence of spina bifida in fetal development). This was also the period in which he launched the notion of antibiotics as growth promoters in animal husbandry. Given the concentration of folic acid in tumours, Jukes also developed the use of antagonists as chemotherapeutic agents, of which the best known (and used still) is methotrexate.

Then came the structure of DNA, in 1953. At a time in life when contemplation of pension entitlements is widely forgiven, Jukes (now back at Berkeley) had re-launched himself by the end of the decade as a molecular biologist. He contributed much to understanding the chemical intricacies of DNA replication, but is best remembered as the first advocate of the doctrine that most genetic mutations are neutral (*vis-à-vis* natural selection).

I first met him in 1960, when he was



Nutritionist turned molecular biologist, and polemicist

camping out in newly occupied quarters in a NASA laboratory on the bay shore on the Oakland side of Berkeley, and I was briefly a science reporter for *The Washington Post*. There was much talk of genetics and a little about the NASA lab's 'mission' (exobiology, I think it was), but much more about the wickedness of the creationists, then flexing their muscles for a fight with the School Board of California over Darwinism's place in the school curriculum.

Jukes had a passionate distaste for cant and, more important, a natural instinct for the effective conduct of a controversy. Excessive politeness, or even bare courtesy, towards opponents might often itself seem like cant. But the crucial determinant of the outcome of a controversy must be a factual foundation that cannot be denied. And then, the controversialist must be assiduous in seeking and using every possible means of making his message heard. Jukes was a past master of the art of formulating questions that his opponents had no hope of answering.

The controversies came thick and fast. Quack cancer cures (such as 'Laetrile', extracted from the inner stones or seeds of peaches) were an early object of his scorn. The rise of the environmental movement in the 1960s forced other targets on his attention, notably the decision of the US government to ban the use of DDT in

agriculture. There was a protracted row about the causes of the decline of the pelican population off the California coast in the early 1970s, which many attributed to the effects of pesticide residues on the thickness of the shells of pelicans' eggs. In the event, DDT was banned, the pelican population recovered — and the controversy was never settled and probably will now never be settled.

In the same spirit, Jukes quickly identified Linus Pauling's advocacy in the 1980s of massive doses of vitamin C as a general prophylactic of human disease, cancers included, as yet another attempt to sell snake-oil to a gullible public, and one that was all the more offensive because of the distinction of its author. For a decade, Jukes was Pauling's most outspoken critic on this score. Not for him the amused indulgence of a great man that marked the research community's general reaction to Pauling's aberration. In the absence of evidence that a snake-oil is effective, Jukes properly held that pretending otherwise is wicked.

Many of these issues were dealt with in Jukes's regular column in *Nature* between 1975 and 1980, but this skeletal version of his career may give the false impression that he was a talented scientist who turned into a cantankerous iconoclast in later life. The issues on which he seized were mostly prompted by his wide experience as a scientist, his knowledge of the importance of naturally occurring food additives (which is what vitamins are), and his deep suspicion that categorical statements of scientific 'fact' are usually exaggerations. His long battle against creationism had different roots — aversion to the misuse of data and the perversion of the style of argument in science. So far as I could tell, Jukes did not relish making enemies although the prospect held no terror for him. But he was a gregarious man, so that losing friends brought a sense of loss.

He was also a cultivated man, who claimed that his admiration for Mark Twain's writing had influenced his decision to move from Canada to the United States. I recall one animated conversation about James Joyce in which Jukes toyed with the argument that Joyce's innovations in the use of language were the equivalent in English literature of the structure of DNA in the life sciences. His conversation was invariably imaginative and a source of great surprise, even fun.

He is survived by his wife Marguerite, an excellent cook, and by three of his four children and by seven grandchildren, who have much to be proud of.

John Maddox

John Maddox is Emeritus Editor of *Nature*.

Precise dating of the Norfolk timber circle

This curious monument on the English coast has survived for more than four millennia.

In August 1998, the site of a subcircular ring of 55 oak timbers surrounding a large inverted oak tree was discovered within the intertidal zone at Holme-next-the-Sea, Norfolk (52.58° N, 00.33° E). The circle, the first surviving example discovered in Britain, was in serious danger of destruction by tidal erosion, so it was imperative that detailed archaeological investigations be undertaken, including dating. Information provided by tree-ring analysis and radiocarbon measurements was combined using a bayesian approach¹, resulting in a precise absolute dating for the structure.

The timber circle (Fig. 1) was dated by taking six samples (four from the ring of posts and duplicate samples from the central tree) for dendrochronological analysis². The ring sequence from the central stump and the four ring sequences from the posts crossmatched and were combined to form a 168-year site chronology.

This tree-ring sequence was compared with a series of reference chronologies. The highest, but statistically non-significant, correlation was against the East Anglia chronology³ ($t=3.98$; higher t -values are more significant⁴), giving an end date for the site chronology of 2050 BC. It also produced lower correlations against East Anglia ending at 2454 BC ($t=3.17$) and 2019 BC ($t=3.14$). Running the ring pattern against the Irish master⁵ gave correlations of $t=3.39$ at 2050 BC, but only $t=0.96$ at 2454 BC and $t=1.7$ at 2019 BC. Thus, although not statistically significant, dendrochronology suggested that 2050 BC is the most likely end date for the site chronology. However, although the consistency of the t -values was encouraging, the correlation values were too low for definitive dating.



Figure 1 The timber circle at Holme-next-the-Sea in Norfolk, England.

For this reason, six contiguous samples of 20 years' growth from the central tree were submitted for high-precision radiocarbon dating. Incorporating the known chronological separation between the samples and the radiocarbon results in a bayesian model (Fig. 2), each of the possible dates suggested by the tree-ring correlations was tested by using the current calibration curve⁶ and OxCal version 2.18 (ref. 7). The hypothesis that the final ring of the tree-ring chronology dates to 2050 BC is consistent with the radiocarbon results ($A=27.7\%$; Fig. 2). However, the radiocarbon evidence is in significant disagreement with the idea that the last ring of the chronology falls in 2454 BC ($A=0.0\%$) or 2019 BC ($A=7.6\%$).

Radiocarbon calibration has undergone a series of refinements in recent years. Calculating this model using other calibration curves shows that the best agreement is produced by using the 1986 calibration curve⁸

($A=46.2\%$), which indicates that inter-laboratory offsets, errors and regional variation in the radiocarbon content of the atmosphere may be significant when producing precise archaeological chronologies⁹.

Combining the information from both dating techniques, including unmeasured partial rings on the outside of the tree-ring samples, reveals that the upturned tree in the centre of the monument died, or was felled, in April to June 2050 BC. The surrounding posts were felled in April to June of the following year, 2049 BC.

Alex Bayliss*, **Cathy Groves†**,
Gerry McCormac‡, **Mike Baillie‡**,
Dave Brown‡, **Mark Brennand§**

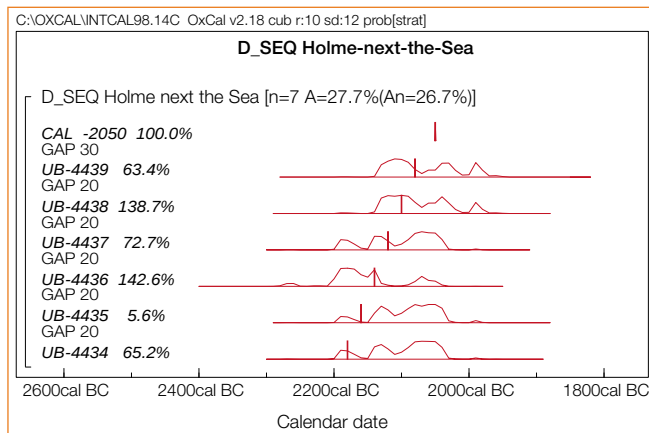
*Ancient Monuments Laboratory,
English Heritage, 23 Savile Row,
London W1X 1AB, UK

†Dendrochronology Laboratory,
Archaeology Research School,
University of Sheffield, Sheffield S1 4DN, UK

‡School of Archaeology and Palaeoecology,
Queen's University, Belfast BT7 1NN, UK

§Norfolk Archaeological Unit,
Garsett House, St Andrew's Hall Plain,
Norwich NR3 1AU, UK
e-mail: alex.bayliss@english-heritage.org.uk

Figure 2 Probability distributions of dates from the central tree bole at Holme-next-the-Sea. Each distribution represents the relative probability that an event occurs at a particular time. For each radiocarbon date, two distributions have been plotted: one in outline (the result of simple radiocarbon calibration), and a solid one (based on the chronological model used). The agreement indices⁷ (A) provide a measure of how well the radiocarbon measurements agree with the absolute date suggested by dendrochronology. The large square bracket on the left and the OxCal keywords define the overall model exactly.



1. Buck, C. E., Cavanagh, W. G. & Litton, C. D. *Bayesian Approach to Interpreting Archaeological Data* (Wiley, Chichester, 1996).
2. English Heritage *Dendrochronology: Guidelines on Producing and Interpreting Dendrochronological Dates* (English Heritage, London, 1998).
3. Baillie, M. G. L. & Brown, D. M. *Br. Archaeol. Rep.* **196**, 543–548 (1988).
4. Baillie, M. G. L. *Tree-ring Dating and Archaeology* 82–85 (Croom Helm, London, 1982).
5. Pilcher, J. R., Baillie, M. G. L., Schmidt, B. & Becker, B. *Nature* **312**, 150–152 (1984).
6. Stuiver, M. *et al. Radiocarbon* **40**, 1041–1083 (1998).
7. Bronk Ramsey, C. *Radiocarbon* **37**, 425–430 (1995).
8. Pearson, G. W. & Stuiver, M. *Radiocarbon* **28**, 839–862 (1986).
9. McCormac, F. G., Baillie, M. G. L., Pilcher, J. R. & Kalin, R. M. *Radiocarbon* **37**, 395–407 (1995).

Transgenic plants

Insecticidal toxin in root exudates from *Bt* corn

Bt corn is corn (*Zea mays*) that has been genetically modified to express insecticidal toxins derived from the bacterium *Bacillus thuringiensis* to kill lepidopteran pests feeding on these plants. Here we show that *Bt* toxin is released into the rhizosphere soil in root exudates from *Bt* corn.

The insecticidal toxin produced by *B. thuringiensis* subsp. *kurstaki* remains active in the soil, where it binds rapidly and tightly to clays¹ and humic acids². The bound toxin retains its insecticidal properties³ and is protected against microbial degradation by being bound to soil particles⁴, persisting in various soils for at least 234 days (the longest time studied), as determined by larvicidal bioassay⁵. Unlike the bacterium, which produces the toxin in a precursor form, *Bt* corn contains an inserted truncated *cry1Ab* gene that encodes the active toxin.

In laboratory studies, caterpillars of the monarch butterfly (*Danaus plexippus*) were killed as a result of feeding on milkweed (*Asclepias curassavica*) that had been artificially contaminated with pollen from transgenic corn that expressed the *cry1Ab* gene from *B. thuringiensis* subsp. *kurstaki*⁶, and green lacewings (*Chrysoperla carnea*), which are insect predators of insect pests, were killed by ingesting European corn borers (*Ostrinia nubilalis*) reared on *Bt* corn⁷.

We germinated surface-sterilized seeds⁸ of *Bt* corn (NK4640Bt) and of the isogenic strain without the *cry1Ab* gene on agar. The seedlings were aseptically placed on plastic screening (6-mm mesh), which, to minimize contamination by products of endosperm hydrolysis, was suspended over 200 ml of Hoagland's solution in 4-litre beakers covered with aluminium foil, which was removed after 18 to 20 days when the tops of the plants had reached it. After 7, 15 and 25 days of growth in a plant-growth room (26 ± 2 °C, 12 h light-dark cycle), the soil-free medium was replaced with fresh solution and analysed.

Total protein in the medium (average of 105 µg protein per plant) was determined by the Lowry procedure⁹. A major band migrating on SDS-PAGE (sodium dodecyl sulphate-polyacrylamide gel electrophoresis) to a position corresponding to a relative molecular mass of 66,000 (66K), the same as that of the *Cry1Ab* protein, was evident after 7 and 15 days only in the exudates from *Bt* corn, although several protein bands of smaller relative molecular mass were seen in exudates from both *Bt* and non-*Bt* corn. We confirmed the presence of the toxin in the exudates from *Bt* corn by immunological assay with Lateral Flow Quickstix (from EnviroLogix, Maine;

detection limit < 10 parts per billion) and verified that it was active in an insecticidal bioassay using larvae of the tobacco hornworm (*Manduca sexta*), a model for testing antilepidopteran activity^{3,5}.

Larvae placed on medium containing exudates from *Bt* corn stopped feeding and began to die after 2 to 3 days and had a mortality of 90 to 95% after 5 days (dose lethal to 50% of larvae, LC₅₀, was 5.2 µg protein). There was no immunological reaction or larval mortality obtained with the exudates from non-*Bt* corn. After 25 days of growth, when the medium was no longer sterile (as demonstrated by streaking it on various microbiological media), the 66K band had disappeared, although there were several new protein bands of smaller relative molecular mass, and the immunological and larvicidal assays were negative, indicating that microbial, and probably also corn, proteases had hydrolysed the toxin.

Samples of soil from the rhizosphere of seedlings that had been transplanted into either sterile or non-sterile soil in test-tubes were taken from randomly selected tubes, vortexed with extraction buffer (EnviroLogix) and centrifuged. We analysed the supernatants using the immunological and larvicidal assays and found that these were positive, even after 25 days of growth, for samples from *Bt* corn (100% mortality; LC₅₀ = 1.6 µg protein per soil tube) but were negative for non-*Bt* corn. Moreover, particles of rhizosphere soil in suspension placed directly on the bioassay medium caused mortality comparable to the supernatants.

Although the concentration of protein in the rhizosphere soil was approximately 95 µg per g soil, the concentration of the actual toxin in the extraction buffer was apparently too low to be detected by SDS-PAGE. These results agree with earlier findings showing that the toxin binds rapidly to surface-active soil particles and that the bound toxin retains its larvicidal activity and is protected by this binding against biodegradation¹⁻⁵.

About 15 million acres of *Bt* corn were planted in the United States in 1998, which was just under 20% of the total acreage of corn¹⁰. The *Bt* toxin that is released into soil from roots during the growth of a *Bt* corn crop would add to the amount of toxin introduced into soil from pollen during tasselling and as a result of the incorporation of plant residues after harvesting the crop.

We have no indication of how soil communities might be affected by *Bt* toxin in root exudates in the field. *Bt* toxin in the rhizosphere might improve the control of insect pests, or it might promote the selection of toxin-resistant target insects. Receptors for the toxin are present in non-target as well as target insects, so there may be a

risk that non-target insects and organisms in higher trophic levels could be affected by the toxin. Further investigations will be necessary to shed light on what might happen underground.

Deepak Saxena*, Saul Flores†, G. Stotzky*

*Laboratory of Microbial Ecology, Department of Biology, New York University, New York, New York 10003, USA
e-mail: gs5@is2.nyu.edu

†Instituto Venezolano de Investigaciones Científicas, Apartado Postal 21827, Caracas 1020A, Venezuela

1. Tapp, H., Calamai, L. & Stotzky, G. *Soil Biol. Biochem.* **26**, 663-679 (1994).
2. Crecchio, C. & Stotzky, G. *Soil Biol. Biochem.* **30**, 463-470 (1998).
3. Tapp, H. & Stotzky, G. *Appl. Environ. Microbiol.* **61**, 1786-1790 (1995).
4. Koskella, J. & Stotzky, G. *Appl. Environ. Microbiol.* **63**, 3561-3568 (1997).
5. Tapp, H. & Stotzky, G. *Soil Biol. Biochem.* **30**, 471-476 (1998).
6. Losey, J. E., Raynor, L. S. & Carter, M. E. *Nature* **399**, 214 (1999).
7. Hilbeck, A., Baumgartner, M., Fried, P. M. & Bigler, F. *Environ. Entomol.* **27**, 480-487 (1998).
8. Benizri, E., Courtade, A. & Guckert, A. *Soil Biol. Biochem.* **27**, 71-77 (1995).
9. Lowry, O. H., Rosebough, N. R., Farr, A. L. & Randall, R. J. *J. Biol. Chem.* **193**, 265-275 (1951).
10. Wadman, M. *Nature* **397**, 636 (1999).

Palaeobiology

Herbivorous diet in an ornithomimid dinosaur

In 1997, twelve well-articulated skeletons of an ornithomimid dinosaur¹ from the Upper Cretaceous Ulansuhai Formation² in China were discovered. Each skeleton contained a preserved gastrolith mass inside the ribcage that was attached on the medial surface of the articulated dorsal ribs and gastralia. The occurrence and characteristics of gastrolith masses in this ornithomimid indicate that these non-avian toothless theropods may have had gizzards and been herbivores, like modern herbivorous birds that use grit to grind up plant matter.

The gastroliths found in these dinosaurs (Fig. 1) are mainly composed of grains of silicate, with no bony elements (energy dispersive X-ray analysis detected hardly any phosphorus pentoxide in the matrix) or insect remains. The gastroliths are found in the same region of each individual, generally close to the middle dorsal vertebrae. One of the isolated gastrolith masses belongs to a large individual (15 × 11 × 3 cm), as judged by the size of its rib impressions, and occupies a greater total volume than those belonging to juveniles (10 × 7.5 × 2.5 cm; Fig. 1c), which is indicative of a body-size correlation to total gastrolith volume like that observed in crocodilians³.

Between species, grit size in modern birds is correlated with body weight, as demonstrated by a regression of mean grain size against the logarithm of mean body weight⁴. We sonically disaggregated part of one gastrolith mass, found in a juvenile, which contained roughly 170 grains per cm³ (grains were counted if they were larger than 0.5 mm along the intermediate axis). The estimated body weight of the juvenile (femur circumference, 57 mm) was about 10 kg (ref. 5), and the mean intermediate diameter of the grain was 1.27 mm, which is small with respect to the estimated body weight and lies outside the range of modern birds (Fig. 1e).

Most gastrolith grains less than 1 mm in diameter vary from angular to very angular,

and those between 1 and 2 mm are mainly subangular (Fig. 1f). Small gastrolith grains may have spent only a short time in the theropod body, whereas larger grains were probably subjected to more abrasion from other grains. Excluding grains ranging from very angular to subangular, the mean grain size is 2.41 mm, which fits the predicted size based on values in modern birds (Fig. 1e). There are over a thousand grains in total, with a concentration of 37 grains per cm³.

The fully terrestrial habitat of this ornithomimid¹ rules out the possibility that the gastroliths were used for hydrostatic adjustment, as has previously been recorded for plesiosaurs⁶, crocodilians³ and a tanga-saurid⁷. The small calcium content of the grains indicates that they were unlikely to

have been used for the uptake of calcium nutrient, as in birds⁸. The limited food-processing ability of the oral cavity of ornithomimids, indicated by a lack of teeth and the weakly developed jaw adductor musculature¹, indicates that the gastroliths may have been used for the mechanical breakdown of food, as in herbivorous dinosaurs⁹ and omnivorous and herbivorous birds⁴.

In modern birds, there is a clear relation between diet and the characteristics of the grit they use⁴: carnivorous birds have no muscular gizzard or grit; frugivorous birds use only a little grit; and herbivorous birds retain more grit than insectivorous or omnivorous birds⁴. The occurrence of gastroliths in this ornithomimid and the large number of grains they contain are consistent with a herbivorous diet and the possession of a muscular ventricular stomach, or gizzard, like that found in modern herbivorous birds.

Gastroliths in non-avian theropods are found in this ornithomimid and in *Caudipteryx*¹⁰. Phylogenetic analysis of Theropoda does not support a close relationship of Ornithomimidae with *Caudipteryx* or the bird clade^{10,11}, indicating that the use of gastroliths in this ornithomimid is the result of convergent evolution, and that herbivory in theropods may have evolved on several occasions.

Yoshitsugu Kobayashi*, Jun-Chang Lu†, Zhi-Ming Dong†, Rinchen Barsbold‡, Yoichi Azuma§, Yukimitsu Tomida¶

*Southern Methodist University, Department of Geological Sciences, Dallas, Texas 75275, USA

†Institute of Vertebrate Paleontology and Paleoanthropology, Academia Sinica, PO Box 643, Beijing 100044, China

‡Geological Institute, Mongolian Academy of Sciences, Ulaan Baatar 11, Mongolia

§Fukui Prefectural Museum, 2-19-15 Omiya, Fukui 910-0016, Japan

¶National Science Museum, 3-23-1 Hyakunincho, Shinjuku, Tokyo 169-0073, Japan

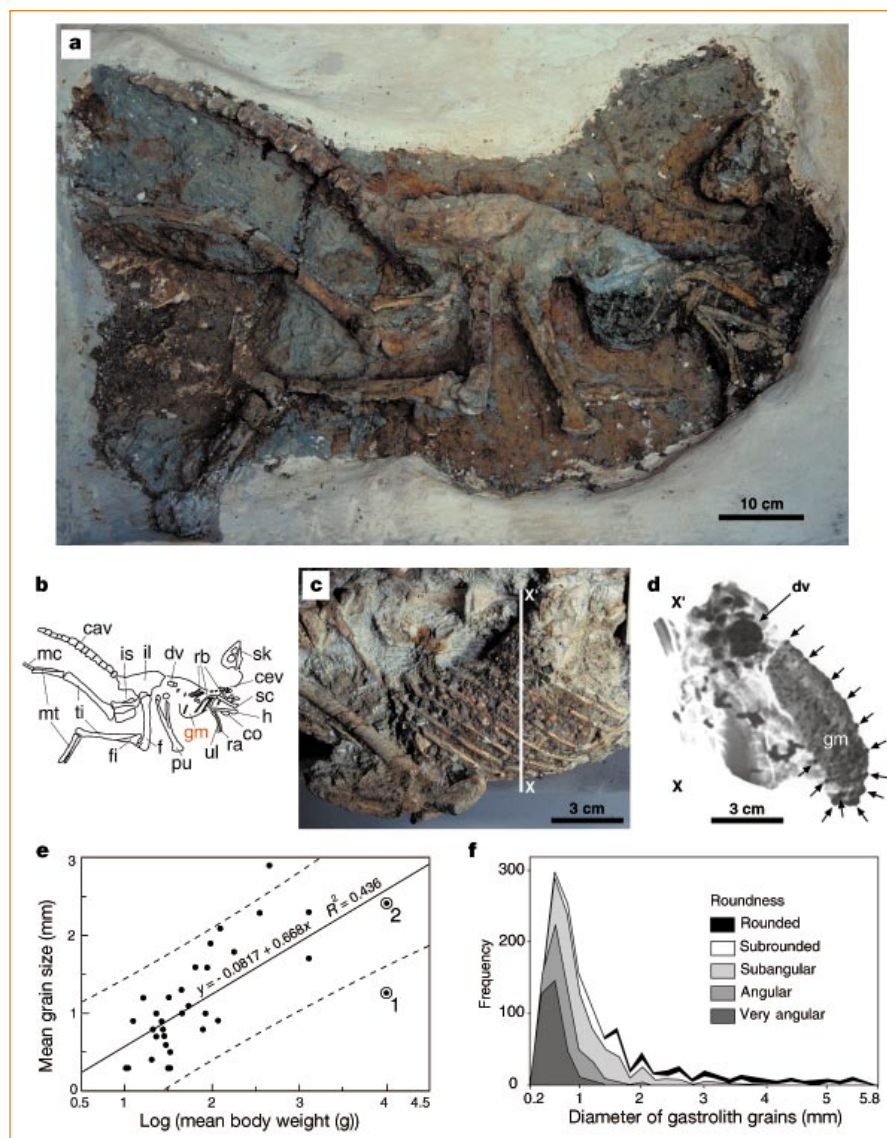


Figure 1 Gastroliths found in the ribcage of skeletons of ornithomimid theropod dinosaurs. **a**, Articulated ornithomimid specimen IVPV-V11797-1. **b**, Schematic representation of the skeleton shown in **a**. **c**, Gastrolith mass in stomach region of IVPV-V11797-4. **d**, Computed tomography scan at the white line (X-X') in **c**. Unlabelled arrows point to ribs. **e**, Scatter plot based on modern birds (single dots)^{4,12} and the ornithomimid (1). Broken lines indicate the 95% confidence interval. The ornithomimid value, excluding very angular to subangular gastrolith grains (2), is shown for comparison. **f**, Grain size frequency subdivided by grain roundness. Abbreviations in **b** as above.

1. Barsbold, R. & Osmólska, H. in *The Dinosauria* (eds Weishampel, D. B., Dodson, P. & Osmólska, H.) 225–244 (Univ. California Press, Berkeley, 1990).
2. Bureau of Geology and Mineral Resources of Nei Mongol Autonomous Region People's Rep. China, Ministry Geol. Miner. Res. Geol. Mem. 25, 1–725 (1991).
3. Cott, H. B. *Trans. Zool. Soc. Lond.* 29, 211–356 (1961).
4. Gionfriddo, J. P. & Best, L. B. *Wilson Bull.* 108, 685–696 (1996).
5. Anderson, J. F., Hall-Martin, A. & Russell, D. A. *J. Zool.* 207, 53–61 (1985).
6. Darby, D. G. & Ojakangas, R. W. *J. Paleontol.* 54, 548–556 (1980).
7. Currie, P. J. *Palaeontol. Afr.* 24, 99–168 (1981).
8. Scott, M. L., Hull, S. J. & Mullenhoff, P. A. *J. Poultry Sci.* 50, 1055–1063 (1971).
9. Farlow, J. O. *Paleobiology* 13, 60–72 (1987).
10. Ji, Q., Currie, P. J., Norell, M. A. & Ji, S. A. *Nature* 393, 753–756 (1998).
11. Xu, X., Wang, X. L. & Wu, X. C. *Nature* 401, 262–266 (1999).
12. Dunning, J. B. Jr in *CRC Handbook of Avian Body Mass* (ed. Dunning, J. B. Jr) 3–310 (CRC Press, Boca Raton, 1993).

Reproduction

Germ cells colonized by endosymbiotic bacteria

Wolbachia (Rickettsiaceae) is a genus of maternally inherited endosymbiotic bacteria commonly found in the reproductive tissues of arthropods. These bacteria manipulate host reproduction to increase the number of infected individuals within the population, erecting intraspecific fertility barriers, causing parthenogenesis or resulting in the feminization of genetic males¹. They are usually transmitted vertically, so we predicted that they should have evolved a mechanism to target the host's germ cells during development. Here we show that *Wolbachia* become concentrated in the germ plasm of the *Drosophila* egg. Mutations in developmental patterning genes² demonstrate that this localization is dependent on the assembly of germ plasm.

The endosymbiotic bacteria (fluorescent staining in Fig. 1) are maternally inherited and sensitive to the antibiotic tetracycline. Analysis of their DNA after amplification by polymerase chain reaction (PCR) revealed an *ftsZ* sequence³ that was identical to that of *Wolbachia* strain A. This sequence was the same in both the Canton S wild-type *Drosophila* stock and the laboratory stock⁴ that we used as a source of endosymbionts for all the other stocks.

These bacteria are found throughout the syncytial embryos of infected strains of *Drosophila*, concentrating in the region of the cortical actin cytoskeleton and accumulating with the mitochondrial motor protein KLP67A on the microtubules of mitotic asters^{4–6}. These mechanisms might ensure transmission of the bacteria, as would colonization of the posterior pole, the site of germ-cell formation. This has been reported for *Nasonia* wasp embryos, which also show cortical localization^{7,8}, but not for *Drosophila*.

To identify the host genes required for posterior localization of *Wolbachia* in the absence of asters, we counted the bacteria in four regions of unfertilized eggs (Table 1) from wild-type females and from females bearing mutations in *gurken* or *oskar*, two maternally acting genes that determine the site of germ-cell formation. The posterior pole region P1 had a mean bacterial density

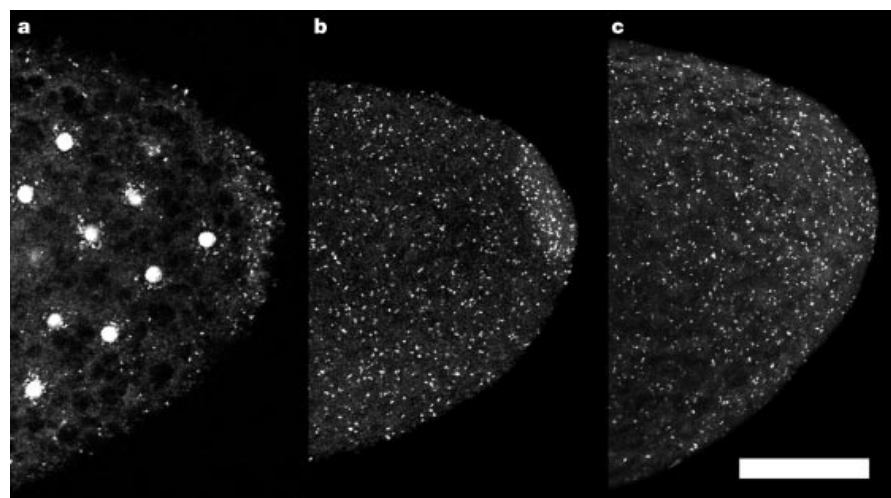


Figure 1 Colonization of *Drosophila* germ cells by *Wolbachia*. **a**, *Wolbachia* clustered on asters around proliferating nuclei and concentrated in the germ plasm on the posterior dorsal side of a wild-type *Drosophila* embryo (single optical section). **b**, Germ-plasm colonization in an unfertilized egg from a heterozygous *oskar*^{5/+} mother. **c**, Unfertilized egg lacking germ-cell determinants from an *oskar*^{5/oskar} mother. Scale bar, 50 μ m. Serial 2- μ m optical sections of *Drosophila* eggs were fixed in methanol and stained for DNA with H33258, collected on a Bio-Rad MRC 1024ES confocal microscope with multi-photon excitation, using a 40 $\times$ 0.85NA (air) lens set at a zoom of 1.5 and a pixel size of 0.388 μ m, and then converted to a single two-dimensional image by maximum brightness projection.

significantly different from those of areas A1, A2 and P2 in eggs from wild-type and heterozygous mothers. The bacteria were not localized in eggs from homozygous mutant mothers. The posterior localization of *Wolbachia* therefore requires not only the functional anteroposterior axis of polarized microtubules specified by Gurken signalling⁹, but also the germ plasm determined by Oskar².

The eggs of our Canton S strain had a density of $3.04 \pm 1.04 \times 10^6$ bacteria per mm³. We used published data⁷ to calculate the density in the *Nasonia vitripennis* egg as $0.51 \pm 0.13 \times 10^6$ bacteria per mm³. In contrast to the tight posterior localization of *Wolbachia* in the *Nasonia* embryo, most bacteria in *Drosophila* colonize somatic cells. The fate of these bacteria is unknown, but they may infect the germ line later in development.

Alternatively, the symbiosis may be less well established in *Drosophila*. We considered that transferring *Wolbachia* from *Nasonia* to *Drosophila* would show whether the germ-cell signals of the two species are recognized by a conserved bacterial mechanism. We found that early germline association seems to be symbiotic: *Wolbachia* infection was not a significant variable influencing the number of germ

cells or the fidelity of their migration to the gonad in the Canton S strain (results not shown).

Posterior localization of *Wolbachia* was also abolished in eggs of *Drosophila* that are mutant for the *vasa* or *tudor* genes, which encode germplasm components localized by Oskar (data not quantified). Functional germ-cell determinants are localized in the mature oocyte (stage 14)¹⁰, but we could find no oocytes at stages 1–14 that had posterior localization of *Wolbachia*, in contrast to 80% of unfertilized eggs 0–3 hours after egg deposition (results not shown).

If germ plasm is merely a rich food source that stimulates local proliferation, why are *Wolbachia* not seen there earlier in oogenesis? Although we have not excluded the possibility of rapid local proliferation after egg deposition, we think it is more likely that the bacteria seen in ovarian nurse cells and oocytes relocate in response to a component of pole plasm that is translated only after egg activation¹¹.

Stephen J. Hadfield, J. Myles Axton

Department of Zoology, University of Oxford,
South Parks Road, Oxford OX1 3PS, UK
e-mail: myles.axton@zoo.ox.ac.uk

Table 1 Quantification of *Wolbachia* in regions of the *Drosophila* egg

Genotype	Bacterial density (mean $\pm$ s.d. $\times 10^6$ per mm ³)			
Canton S	P2(2.57 $\pm$ 1.47)	A2(2.79 $\pm$ 0.98)	A1(3.29 $\pm$ 1.10)	
P1(4.65 $\pm$ 2.10)				
<i>gurken</i> ^{2/+}	A2(2.04 $\pm$ 0.94)	A1(2.19 $\pm$ 0.76)	P2(2.34 $\pm$ 1.17)	P1(3.99 $\pm$ 1.65)
<i>oskar</i> ^{5/+}	P2(2.22 $\pm$ 1.27)	A2(2.94 $\pm$ 1.67)	A1(2.96 $\pm$ 1.57)	
P1(4.26 $\pm$ 1.80)				
<i>gurken</i> ^{2/gurken}	A2(2.10 $\pm$ 1.32)	P2(2.24 $\pm$ 1.00)	A1(2.29 $\pm$ 0.89)	P1(2.45 $\pm$ 1.36)
<i>oskar</i> ^{5/oskar}	P2(1.88 $\pm$ 0.66)	P1(1.95 $\pm$ 0.66)	A2(2.06 $\pm$ 0.56)	A1(2.11 $\pm$ 0.76)

Bacterial density in 10 unfertilized eggs of each genotype, measured in a 95 $\times$ 29 μ m ellipse placed at the anterior pole (A1), two sites (A2 and P2) at

1. Werren, J. H. *Annu. Rev. Entomol.* **42**, 587–609 (1997).
2. Rongo, C. & Lehmann, R. *Trends Genet.* **12**, 102–109 (1996).
3. Holden, P. R., Brookfield, J. F. Y. & Jones, P. *Mol. Gen. Genet.* **240**, 213–220 (1993).
4. Glover, D. M. *et al. Nature* **348**, 117–117 (1990).
5. Callaini, G., Riparbelli, M. G. & Dallai, R. *J. Cell Sci.* **107**, 673–682 (1994).
6. Pereira, A. J., Dalby, B., Stewart, R. J., Doherty, S. J. & Goldstein, L. S. B. *J. Cell Biol.* **136**, 1081–1090 (1997).
7. Breeuwer, J. A. J. & Werren, J. H. *Nature* **346**, 558–560 (1990).
8. Breeuwer, J. A. J. & Werren, J. H. *Genetics* **135**, 565–574 (1993).
9. Gonzalez-Reyes, A., Elliot, H. & St Johnston, D. *Nature* **375**, 654–658 (1995).
10. Illmensee, K. & Mahowald, A. P. *Proc. Natl Acad. Sci. USA* **71**, 1016–1020 (1974).
11. Jongens, T. A., Hay, B., Jan, L. Y. & Jan, Y. N. *Cell* **70**, 569–584 (1992).

Synexpression groups in eukaryotes

Christof Niehrs & Nicolas Pollet

Division of Molecular Embryology, Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 280, D-69120 Heidelberg, Germany

In 1960, Jacob and Monod described the bacterial operon, a cluster of functionally interacting genes whose expression is tightly coordinated. Global expression analysis has shown that the highly coordinate expression of genes functioning in common processes is also a widespread phenomenon in eukaryotes. These sets of co-regulated genes, or 'synexpression groups', show a striking parallel to the operon, and may be a key determinant facilitating evolutionary change leading to animal diversity.

Differential gene activity largely accounts for the regulation of the genetic program underlying cell differentiation and metazoan development. Turning genes on and off at the right time determines cell fate, metabolic state and division. The development of DNA micro-arrays for the simultaneous monitoring of thousands of transcripts constitutes a breakthrough that has allowed global insights into gene expression. Now that the bigger picture is emerging, the expression data reveal a high degree of order in the genetic program: a tight coordination of the expression of groups of genes functioning in a common process. This phenomenon was initially discovered 40 years ago in the bacterial operon, but the concept was dismissed for eukaryotes because of their different mode of gene regulation.

Here we review groups of co-regulated eukaryotic genes and discuss their functional significance and relationship to other aspects of the modular genetic architecture. We focus on the evolutionary implications of coordinate gene expression because excellent reviews on application and results¹⁻³, methodological⁴⁻⁶ and bioinformatic⁷ aspects of DNA micro-arrays are available.

Identification of synexpression groups

In a typical gene-expression monitoring experiment using DNA micro-arrays, transcript levels are measured in cells after experimental treatment at various time points to yield an expression profile. Where such analyses have been carried out with sufficient temporal resolution, a correlation between the expression profiles of genes functioning in the same process has been found. In yeast, where expression of the entire genome can be monitored, 17 genes were induced by overexpression of Yap-1, a transcription factor conferring resistance to various noxious agents⁸. Of these, 9 genes encode different kinds of oxido-reductases that are thought to act in detoxification. The transcriptional program of yeast after sporulation⁹ and during cell cycle^{10,11} has been analysed by the same technique. In these studies, distinct temporal expression profiles were observed, and co-expression of genes involved in a common process, such as replication, chromosome pairing and mitosis, was a recurring theme (Table 1).

The response of starved human fibroblasts to serum addition has also been followed over several hours. A striking observation was not only the induction of genes involved in wound healing and the G2/M transition, but in particular the coordinate expression of five genes involved in cholesterol biosynthesis (Figs 1a, 2a)¹². Groups of co-regulated genes were also identified by producing a high-resolution temporal expression profile during spinal cord development¹³. Using polymerase chain reaction with reverse transcription (RT-PCR) analysis, five basic waves of expression were discovered, and a number of genes acting in GABA (γ -aminobutyric acid) signalling mapped to one particular expression profile (Figs 1b, 2b).

Although this suggests that the co-expression of genes involved in the same biological process is a widespread phenomenon, in many expression profiling studies, the functional correlation of member genes is often very imperfect because of noise from false-positive

genes that happen to be co-regulated. For example, the immediate early genes *Krox20* and *junB* were co-induced in a global-expression profiling experiment after growth factor treatment¹⁴; however, these genes have very different expression patterns and function during early mouse embryogenesis. This 'noise' is due to both the limited number of samples constituting the expression profiles, and the fact that expression profiles are typically clustered under one specific experimental condition, for example, diauxic shift, sporulation or serum addition, which each addresses only one sector of the regulatory repertoire of a cell.

Two global expression analyses have reported a very tight correlation between function and expression^{15,16}. In one, a cluster analysis was carried out on all yeast genes, aggregating data of about 80 experiments from diauxic shift, cell cycle, sporulation, temperature and reducing shocks. Under these diverse conditions, 10 clusters of an exceptionally tight correlation between expression profile and function was apparent¹⁵ (Table 1). The largest cluster is formed by more than 100 genes involved in protein synthesis —either ribosomal proteins or translation factors. Tight correlation was also observed for genes encoding components of large molecular complexes such as the proteasome or chromatin, and for genes encoding components of the glycolytic pathway (Table 1).

We used *Xenopus* embryos for expression monitoring by whole mount *in situ* hybridization to identify differentially expressed genes. Of the 273 genes identified, 38 could be grouped into four gene clusters that each share a very distinctive, complex expression pattern¹⁶. An example is the growth factor bone morphogenetic protein 4 (BMP-4) group (Figs 1c, 2c), members of which are expressed dorsally in the eye, heart, tailbud and lateral plate mesoderm of tailbud-stage *Xenopus* embryos. This group consists of seven members which all encode components of the BMP signalling pathway, as studied in early dorso-ventral patterning of mesoderm (Fig. 2c).

The largest group (25 members) identified in *Xenopus* is the chromatin group. These genes are characterized by their repression in tissues becoming postmitotic; and most encode chromatin proteins (for example, histones or HMG proteins), or proteins indirectly interacting with chromatin such as ornithin decarboxylase, a key enzyme in spermidin synthesis (Table 1).

We proposed the term 'synexpression groups' to designate sets of genes that share a complex 'spatial' expression pattern (multiple tissues), and that function in the same process^{16,17}. However, we propose that the term could be extended to groups of genes that share complex 'temporal' (multiple time points) or 'experimental' (multiple experimental conditions) expression profiles, because the currently used term 'gene cluster' is ill defined: traditionally, it refers to physically linked genes and is therefore misleading in the context of co-expression. Also, synexpression groups allow the prediction of gene function with high precision, unlike many noise-laden co-expressed gene clusters, and they probably confer emergent properties during evolution (see below).

Notably, synexpression refers to groups of genes co-expressed

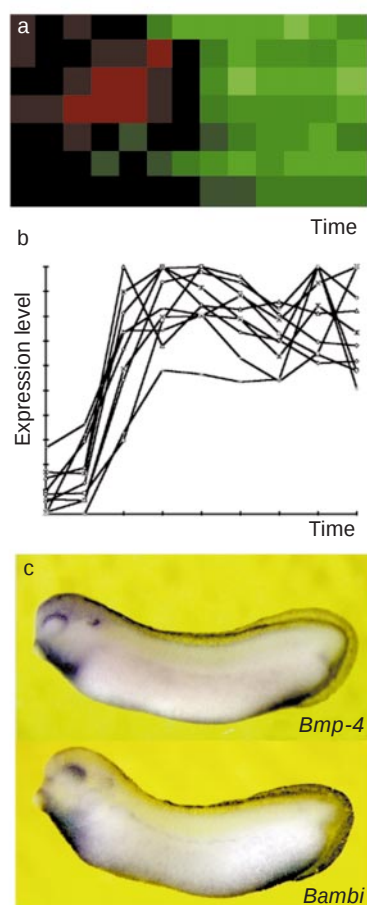


Figure 1 Identification of synexpression groups. **a**, Cholesterol biosynthesis group. Cluster image resulting from large-scale gene expression analysis using DNA micro-arrays in the response of human fibroblasts to serum. Each line depicts the temporal expression pattern of a gene. HMG CoA reductase (1); IPP isomerase (triplicate) (2–4); farnesyl-diphosphate farnesyl transferase (5); squalene epoxidase (6); cytochrome P450 lanosterol 14- α -demethylase (7). Red levels, fold repression; black levels, no change; green levels, fold induction¹². **b**, RT-PCR analysis of genes expressed in developing rat spinal cord reveals a group of co-expressed genes defining a GABA synexpression group¹³. **c**, Wholemount *in situ* hybridization of BMP-4 and BAMBI in *Xenopus* tailbud embryos, showing the expression of members of the BMP-4 synexpression group¹⁶. Expression patterns are highly similar, for example in dorsal eye, heart, proctodeum and lateral plate mesoderm.

under diverse conditions and hence their identification typically requires expression profiling at high resolution. This is naturally the case with *in situ* hybridization, in which the expression of one gene is profiled over thousands of samples (that is, different cells). However, screening by *in situ* hybridization is only practical for a few hundred transcripts per week. The advantage of DNA micro-arrays is that the experiment is carried out with thousands of genes simultaneously; the drawback is the limited number of samples constituting the expression profile (typically less than 20). In practical terms, it is not useful to consider genes induced in one or a few samples as a synexpression group, they may be co-regulated only under these conditions¹⁴. Similarly, muscle-specific genes are not necessarily members of the same synexpression group, as these genes may be involved in diverse, yet muscle-specific biological processes (for example, respiration, contraction and cell adhesion). A gene that is member of one synexpression group cannot be part of another.

In summary, tight spatio-temporal co-expression of genes functioning in the same cell-biological process is a widespread phenomenon in eukaryotes. The proteins encoded by genes of synexpression

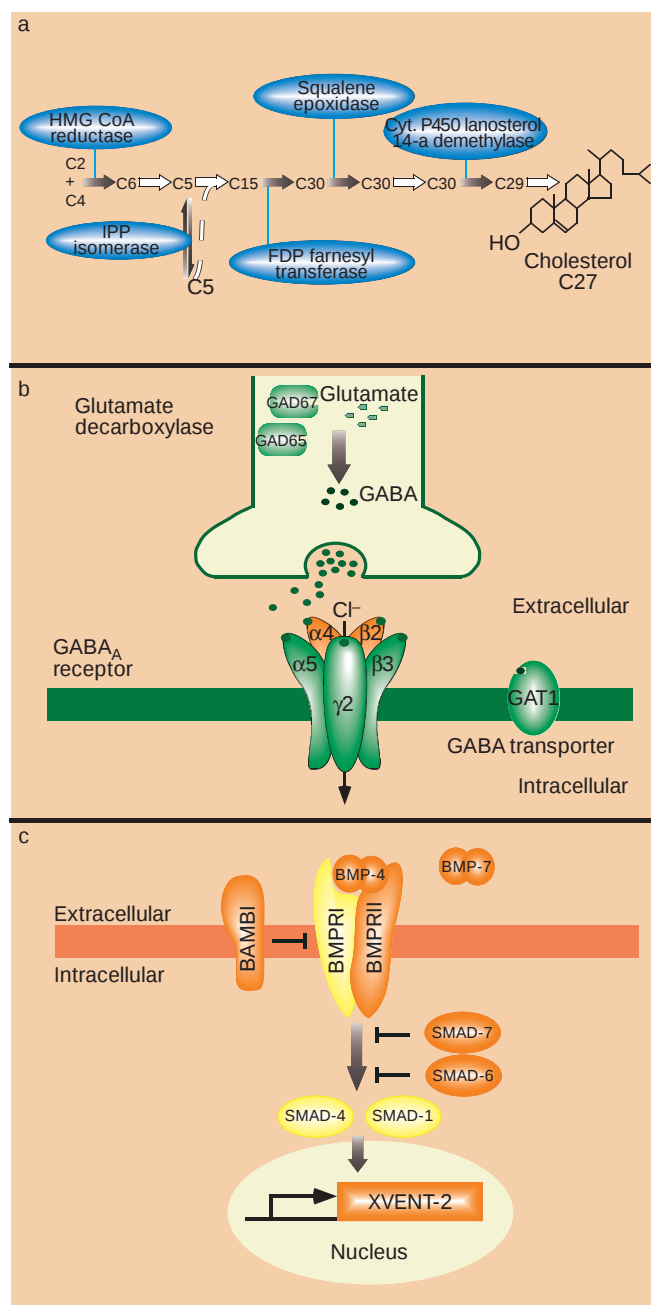


Figure 2 Synexpression groups. **a**, The cholesterol biosynthesis synexpression group and pathway, with co-expressed enzymes (see Fig. 1a) indicated (blue). C(*n*), carbon backbones. Filled and empty arrows, one or multiple enzymatic reactions, respectively. Data from ref. 12. **b**, The GABA synexpression group, with co-expressed components of GABA signal transmission (see Fig. 1b) indicated (green). Glutamate decarboxylase (isoforms GAD65 and GAD67) catalyses the decarboxylation of glutamate into the inhibitory neurotransmitter GABA, which binds to the GABA_A/benzodiazepine receptor, a ligand-gated chloride channel. The pentameric GABA_A receptor is composed of multiple chain subtypes. GAT1 terminates the action of GABA by re-uptake into presynaptic terminals. For simplicity, not all GABA_A receptor subtypes are shown. Data from ref. 13. **c**, The BMP-4 synexpression group and signalling pathway, with co-expressed components of the BMP signal transduction pathway (Fig. 1c) indicated (orange). BMP-4 and BMP-7 are ligands for type I and type II BMP transmembrane receptors. Upon ligand binding, type II receptors phosphorylate SMAD-1 transcriptional activators which complex with SMAD-4 and enter the nucleus, activating transcription of target genes such as XVENT-2. Inhibitory SMADs such as SMAD-6 and SMAD-7, as well as BAMBI⁴¹, interfere with signal transduction. Data from refs 16 and 38–41.

Table 1 Examples of synexpression groups

Synexpression group (model system)	No. of genes	Method	Significance	References
DNA structure				
Histone (yeast)	8, 9	DNA chip	Chromatin structure	10, 15, 11
Chromatin (<i>Xenopus</i>)	25	<i>In situ</i> hybridization	Cell cycle regulated DNA scaffold components	16
Protein biosynthesis				
Ribosome and translation (yeast)	112	DNA chip	Protein biosynthesis	15
ER-import (<i>Xenopus</i>)	7	<i>In situ</i> hybridization	Components of ER/secretion machinery	16
mRNA splicing (yeast)	14	DNA chip	Gene expression	15
Cell cycle				
Spindle pole body assembly (yeast)	11	DNA chip	Cell division	15
Early (I) induction (yeast)	62	DNA chip	Sporulation, entry in meiosis	9
MCM complex (yeast)	34	DNA chip	DNA replication	15, 11
CIB2 (yeast)	35	DNA chip	Mitosis	11
CLN2 (yeast)	76	DNA chip	DNA replication	11
MAT (yeast)	13	DNA chip	Mating	11
G ₂ /M transition (yeast)	9	DNA chip	DNA replication	12
Mitochondrial function				
Mitochondrial ribosome (yeast)	22	DNA chip	Mitochondrial protein biosynthesis	15
Mitochondrial ATP Synthase (yeast)	15	DNA chip	ATP Synthase	15
Mitochondrial TCA cycle (yeast)	16	DNA chip	Mitochondrial metabolism	15
Metabolism				
Glycolysis (yeast)	17	DNA chip	Carbon metabolism	15
Methionine biosynthesis (yeast)	20	DNA chip	Sulfur amino acid metabolism	11
Cholesterol biosynthesis (human fibroblasts)	5	DNA chip	Membrane biosynthesis	12
Cell signalling				
BMP-4 (<i>Xenopus</i>)	7	<i>In situ</i> hybridization	BMP signalling	38,39,16,40
Delta1 (<i>Xenopus</i>)	5	<i>In situ</i> hybridization	Delta-Notch signalling	16
GABA (rat)	9	RT-PCR	GABA neurotransmission	13
Others				
Yap-1 response (yeast)	17	DNA chip	Resistance to oxidative stress	8
Proteasome (yeast)	27	DNA chip	Protein degradation	15

Groups of coordinately expressed genes with tight correlation between expression profiles and role of the genes in a defined process were selected. ER, endoplasmic reticulum. TCA, tricarboxylic acid.

groups are not necessarily homologous (for example, receptor, ligand and transcription factor in the BMP-4 synexpression group; Fig. 2c) and hence co-expression is not the consequence of gene duplication. The close correlation between expression pattern and gene function in synexpression implies that high-resolution expression profiling allows the assembly of genes into groups that define molecular pathways, and allows predictions to be made about the roles of constituent members with unknown function.

Such prognoses have already proved successful^{9,18,19}, and although the groups of genes did not reveal a genetic hierarchy nor a biochemical protein function, they permitted the formulation of specific and testable hypotheses which were rapidly confirmed experimentally. A similar route is being taken in microbial genetics, in which DNA sequencing of entire bacterial genomes has identified new operons and unknown gene functions using this 'guilt-by-association' paradigm²⁰. Furthermore, as at least three synexpression groups contribute to mitochondrial function (ATP synthesis, trichloroacetic acid cycle and mitochondrial ribosome¹⁵; Table 1), synexpression groups may reveal subroutines in biological programmes.

Synexpression groups versus operons

The first example of tight integration of gene expression and function was the *lac* operon, a transcriptional unit of genes functioning in a common metabolic pathway in bacteria utilizing lactose²¹. The genes encode *lacZ*, a hydrolase cleaving lactose, *lacY*, a galactoside permease and *lacA*, a galactoside transacetylase. The co-expression of these genes is achieved by their organization in a linear transcription unit that is regulated by a repressor. In prokaryotes, genes of many pathways are organized into operons, and remarkably there appear to be operons for cell division and protein export^{20,22}, similar to the respective yeast and *Xenopus* synexpression groups (Table 1). Estimates are that only a minority of genes is organized in operons (13%)²³ or synexpression groups (5–10%)^{15,16}.

Operons are not generally found in eukaryotes. Genes are not transcribed as polycistronic messenger RNA but are individually regulated by complex promoters and are arranged on chromosomes without correlation regarding their function or expression, although there are exceptions²⁴. Hence, co-expression is unlikely to be achieved by *cis*-regulation of physically associated transcription units, as many genes of a synexpression group are located on different chromosomes. Synexpression is more probably coordinated by one or a few *trans*-acting factors regulating common promoter elements. Such promoter elements and their candidate transcriptional regulators have been identified in yeast^{8,9,11}. For example, a single ECB *cis*-regulatory element is sufficient to confer early G1 expression of *cdc6*, *cdc46* and *swi4* through the MCM1 *trans*-acting factor²⁵. This simple regulatory mode contrasts with the majority of higher eukaryotic genes with complex promoters that are controlled by numerous *trans*-acting factors²⁶.

In the operon, the *lac* repressor serves simultaneously as the transcriptional regulator and as a 'sensor'²⁷ that integrates cellular signals, in this case the availability of lactose, to orchestrate expression of a battery of structural genes. Although such simple modes may also operate in synexpression groups, it can be envisaged that the site of integration of complex signals is made upstream of transcriptional regulators, for example, through growth factors.

Evolutionary implications

The concept of synexpression groups implies that within individual member genes, there is co-evolution of function and expression, in other words, co-evolution of promoter and coding sequences. Hence, there must be selective pressure promoting this type of genetic organization. There are two modes that may explain the adaptability of synexpression groups, one benefiting the individual organism, the other its capacity to evolve. We will only outline the possible consequences of modular gene expression on evolution.

For the individual organism, operons and synexpression groups

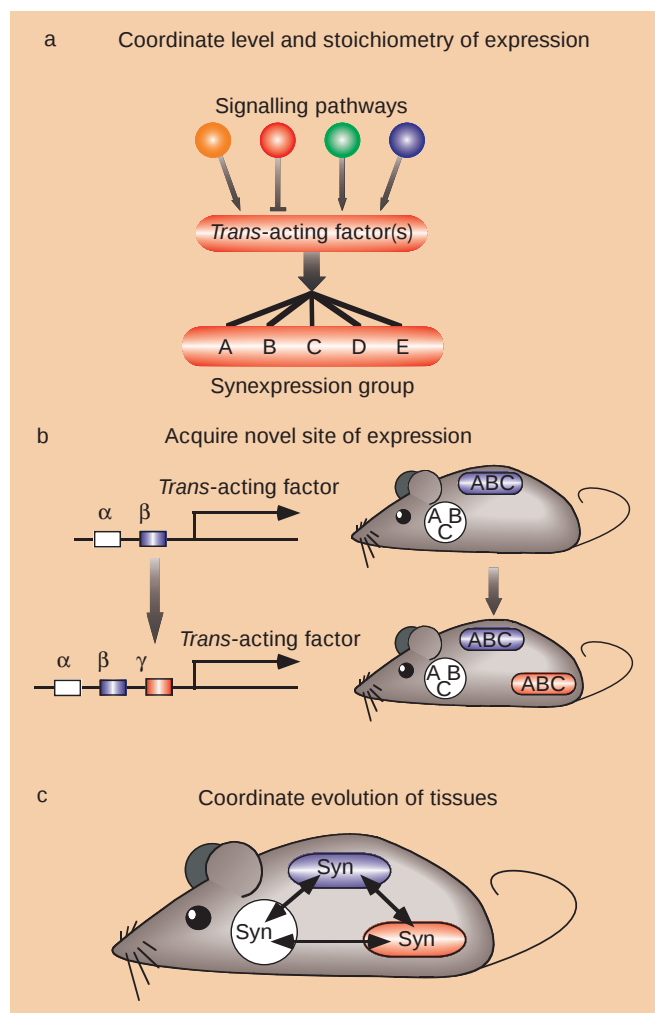


Figure 3 Adaptive advantages of synexpression groups. **a**, A battery of interacting genes (A–E) is co-controlled through common *cis*-regulatory elements by a *trans*-acting factor (or a network thereof) that is regulated by inputs from various signalling pathways. Common regulation of expression of functionally interacting genes allows co-regulation of expression level, for example, upon the appropriate stimulus, and promotes synthesis of individual gene products in stoichiometric amounts. More complex interactions between individual members of synexpression groups (cross-talk and feedback mechanisms) are possible but are not shown for simplicity. **b**, A battery of interacting genes (A–C) is co-regulated by a *trans*-acting factor, which in turn is controlled by tissue- (or stage-) specific promoter elements, α and β . Acquisition of promoter element, γ , allows deployment of the entire synexpression group in a new site. Note that this enhanced integration is paralleled by an increased genetic constraint due to the coupling of genetic modules. Loss of the individual elements α , β or γ may lead to rudimentation of the respective organ. Synexpression groups may regulate each other leading to highly complex gene networks. **c**, Tissues sharing synexpression groups (Syn) will experience evolutionary covariation, which leads to physiological integration and to co-evolution of the linked tissues. The consequences are enhanced rate of evolution and increased pleiotropy and genetic constraints (see refs 33 and 42).

may be adaptive because they allow the co-expression of a set of genes at high level only when required, for example, the *lac* operon in the presence of galactose or the detoxification group in the presence of poisons. Apart from energetic economy, interacting gene products frequently need to assemble stoichiometrically or may require co-translation for forming a complex²⁰, which is promoted by co-expression (Fig. 3a). Therefore, components of supramolecular complexes will probably be organized in synexpres-

sion groups, as is the case for the proteasome, ribosomal machinery and the chromatin complex (Table 1).

The other major reason why synexpression groups may be adaptive is because they enhance the modularity of the genetic potential and promote rapid evolution. The primary source leading to animal diversity is not thought to be the differences in gene products but the differences in the networks in which the gene products are connected during development. The regulatory linkages that make the function of gene products or whole genetic modules contingent on extracellular conditions are thought to be a major factor facilitating evolutionary change^{27,28–30}. The concept of such genetic modules, also called syntags or gene cassettes, was developed by *Drosophila* geneticists when it was realized that embryogenesis can be broken down into discrete elementary operations, such as segmentation, neurogenesis or appendage formation. These processes often encompass the same genetic pathways at multiple stages and multiple places of development^{31,30}, for example, the mitogen-activated protein kinase pathway used in the development of *Drosophila* eye, *Caenorhabditis elegans* vulva and *Xenopus* mesoderm³². By definition, every synexpression group is a gene cassette or a subset thereof.

Compartmentalizing an organism's genetic potential in synexpression groups may produce evolutionary change requiring few mutational steps. In bacteria, for example, the horizontal mode of gene transfer may drive the evolution of operons: during conjugation only parts of the genome may be exchanged, and thus physical proximity of interacting genes promotes transfer of complete functional units²². A similar reasoning applies to eukaryotes and metazoa. First, through acquisition or loss of a single region-specific promoter element controlling relevant *trans*-acting factors, an entire synexpression battery may be re-deployed in a new developmental context (time or site of expression) or be connected to other regulatory pathways (Fig. 3b). An enormous complexity of regulatory networks can be built with this model, assuming that higher order *trans*-acting factors integrate multiple synexpression groups. In the classic Britten–Davidson model of coordinate gene regulation, these and other adaptive features of integrated gene expression have been considered, although the proposed mechanism for co-expression involving regulatory antisense RNAs has not been borne out²⁷.

Furthermore, it may be adaptive to tie certain tissues sharing the same synexpression group genetically together because it allows co-evolution of those tissues, including properties such as size, shape and enzyme composition³ (Fig. 3c). For example, the precursors of fore- and hindlimbs were genetically linked together early in their history by the common expression of *Hox* genes. This may explain why digits arose at the same time in hand and foot; these appendages could evolve together presumably because their development was brought under the control of a common enhancer for the *Hox* complex³⁴. However, although it is reasonable that the genetic linkage of appendages may lead to adaptive co-diversification of these functionally related structures, it is less obvious which properties are shared by the central nervous system and developing somites that makes it advantageous to have them co-controlled by the Delta–Notch pathway¹⁶ (Table 1). Possibly, it is the need to select one or a few cells from an equivalent group to take on one particular cell fate, which is promoted by the lateral inhibition mediated by the Notch pathway³⁵. However, a combination of tissues expressing a synexpression group may just reflect evolutionary history—the chance acquisition of a functional module proving to be adaptive.

Only a subset of functionally interacting genes is found in synexpression groups. For example, some but not all genes involved in methionine biosynthesis are part of the respective synexpression group¹¹, and *SMAD-1* and *SMAD-4* which also participate in BMP-4 signalling are ubiquitously expressed during early embryogenesis^{36,37} and thus are not part of the BMP-4 group. One possible explanation for such non-selected genes is that they are

evolutionarily younger and may not have had enough time to be recruited. Another reason may be that non-recruited components serve in multiple processes (for example, SMAD-4 in BMP, TGF- β , as well as activin pathways) and are not tightly coupled to any one synexpression group. This suggests that a gene involved in multiple pathways may rarely be a member of a synexpression group.

Once a gene is recruited into a synexpression group it should be very stable because of the connectivity in multiple tissues, and we propose that members of synexpression groups will be conserved even between different animal phyla. Indeed, there are chromatin synexpression groups both in yeast and in *Xenopus* (Table 1); and in the homologous *Drosophila* dpp and *Xenopus* Bmp-4 pathways, inhibitory SMADs (*dad*; SMAD6/7) are co-expressed with their ligands. In contrast, the region-specific deployment of individual synexpression groups by changing the regulation of a sensor²⁷, may be subject to much greater variability because it requires few genetic changes.

Conclusion

Synexpression groups reflect the functional compartmentalization of the eukaryote genome and have a striking parallel to the prokaryote operon. DNA micro-array expression analysis is clearly the most efficient way to identify more synexpression groups, but a future challenge will be to carry out profiling over hundreds or thousands of conditions to map the genomic programme at high resolution. We will probably encounter synexpression groups that make 'no sense' at first glance, in other words, groups in which the functions of the co-expressed genes have no common denominator. These might be particularly interesting, as they might reveal an otherwise hidden logic of cellular regulation. Where synexpression groups couple seemingly unrelated tissues, we might search for adaptive reasons for their integration. The hardwired correspondence between gene function and regulation shown by synexpression groups unveils a degree of order that was previously unsuspected and provides us with new opportunities to study genetic networks. □

1. Brown, P. O. & Botstein, D. Exploring the new world of the genome with DNA microarrays. *Nat. Genet.* **21**, 33–37 (1999).
2. Gerhold, D., Rushmore, T. & Caskey, C. T. DNA chips: promising toys have become powerful tools. *Trends Biochem. Sci.* **24**, 168–173 (1999).
3. Lipshutz, R. J., Fodor, S. P., Gingeras, T. R. & Lockhart, D. J. High density synthetic oligonucleotide arrays. *Nat. Genet.* **21**, 20–24 (1999).
4. Bowtell, D. D. Options available—from start to finish—for obtaining expression data by microarray. *Nat. Genet.* **21**, 25–32 (1999).
5. Cheung, V. G. *et al.* Making and reading microarrays. *Nat. Genet.* **21**, 15–19 (1999).
6. Duggan, D. J., Bittner, M., Chen, Y., Meltzer, P. & Trent, J. M. Expression profiling using cDNA microarrays. *Nat. Genet.* **21**, 10–14 (1999).
7. Bassett, D. E., Eisen, M. B. & Boguski, M. S. Gene expression informatics—it's all in your mine. *Nat. Genet.* **21**, 51–55 (1999).
8. DeRisi, J. L., Vishwanath, R. I. & Brown, P. O. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* **278**, 680–686 (1997).
9. Chu, S. *et al.* The transcriptional program of sporulation in budding yeast. *Science* **282**, 699–705 (1998).
10. Cho, R. J. *et al.* A genome-wide transcriptional analysis of the mitotic cell cycle. *Mol. Cell* **2**, 65–73 (1998).

11. Spellman, P. T. *et al.* Comprehensive identification of cell cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization. *Mol. Biol. Cell* **9**, 3273–97 (1998).
12. Vishwanath, R. I. *et al.* The transcriptional program in the response of human fibroblasts to serum. *Science* **283**, 83–87 (1999).
13. Wen, X. *et al.* Large-scale temporal gene expression mapping of central nervous system development. *Proc. Natl Acad. Sci. USA* **95**, 334–339 (1998).
14. Fambrough, D., McClure, K., Kazlauskas, A. & Lander, E. S. Diverse signaling pathways activated by growth factor receptors induce broadly overlapping, rather than independent, sets of genes. *Cell* **97**, 727–741 (1999).
15. Eisen, M. B., Spellman, P. T., Brown, P. O. & Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl Acad. Sci. USA* **95**, 14863–14868 (1998).
16. Gawantka, V. *et al.* Gene expression screening in *Xenopus* identifies molecular pathways, predicts gene function and provides a global view of embryonic patterning. *Mech. Dev.* **77**, 95–141 (1998).
17. Niehrs, C. Gene-expression screens in vertebrate embryos: more than meets the eye. *Genes Funct.* **1**, 229–231 (1997).
18. Onichtchouk, D. *et al.* The Xvent-2 homeobox gene is part of the BMP-4 signaling pathway controlling dorso-ventral patterning of *Xenopus* mesoderm. *Development* **122**, 3045–3053 (1996).
19. Jen, W. -C., Gawantka, V., Pollet, N., Niehrs, C. & Kintner, C. Periodic repression of Notch pathway genes governs the segmentation of *Xenopus* embryos. *Genes Dev.* **13**, 1486–1499 (1999).
20. Dandekar, T., Snel, B., Huynen, M. & Bork, P. Conservation of gene order: a fingerprint of proteins that physically interact. *Trends Biochem. Sci.* **273**, 324–328 (1998).
21. Jacob, F. The operon—25 years later. *C. R. Acad. Sci. Paris* **320**, 199–206 (1997).
22. Lawrence, J. G. Selfish operons and speciation by gene transfer. *Trends Microbiol.* **5**, 355–359 (1997).
23. Blattner, F. R. *et al.* The complete genome sequence of *Escherichia coli* K-12. *Science* **277**, 1453–1474 (1997).
24. Blumenthal, T. Gene clusters and polycistronic transcription in eukaryotes. *BioEssays* **20**, 480–487 (1998).
25. McInerney, C. J., Partridge, J. F., Mikesell, G. E., Creemer, D. P. & Breeden, L. L. A novel Mcm1-dependent element in the SWI4, CLN3, CDC6, and CDC47 promoters activates M/G1-specific transcription. *Genes Dev.* **11**, 1277–1288 (1997).
26. Yuh, C. H., Bolouri, H. & Davidson, E. H. Genomic *cis*-regulatory logic: experimental and computational analysis of a sea urchin gene. *Science* **279**, 1896–1902 (1998).
27. Britten, R. J. & Davidson, E. H. Gene regulation for higher cells: a theory. *Science* **165**, 349–357 (1969).
28. Gerhart, J. & Kirschner, M. *Cells, Embryos and Evolution* (Blackwell, Malden, 1997).
29. Duboule, D. & Wilkins, A. S. The evolution of 'bricolage'. *Trends Genet.* **14**, 54–59 (1998).
30. Huang, F. Syntags in development and evolution. *Int. J. Dev. Biol.* **42**, 487–494 (1998).
31. Jan, Y. N. & Jan, L. Y. Functional gene cassettes in development. *Proc. Natl Acad. Sci. USA* **90**, 8305–8307 (1993).
32. Nishida, E. & Gotoh, Y. The MAP kinase cascade is essential for diverse signal transduction pathways. *Trends Biochem. Sci.* **18**, 128–131 (1993).
33. Cheverud, J. M. Developmental integration and the evolution of pleiotropy. *Am. Zool.* **36**, 44–50 (1996).
34. Shubin, N., Tabin, C. & Carroll, S. Fossils, genes and the evolution of animal limbs. *Nature* **388**, 639–648 (1997).
35. Artavanis-Tsakonas, S., Rand, M. D. & Lake, R. J. Notch signaling: cell fate control and signal integration in development. *Science* **284**, 770–776 (1999).
36. Lagna, G., Hata, A., Hemmati-Brivanlou, A. & Massague, J. Partnership between DPC4 and SMAD proteins in TGF- β signalling pathways. *Nature* **383**, 832–836 (1996).
37. Meersman, G. *et al.* The C-terminal domain of Mad-like signal transducers is sufficient for biological activity in vivo and transcriptional activation. *Mech. Develop.* **61**, 127–140 (1997).
38. Bhushan, A., Chen, Y. & Vale, W. SMAD7 inhibits mesoderm formation and promotes neural cell fate in *Xenopus* embryos. *Dev. Biol.* **200**, 260–268 (1998).
39. Frisch, A. & Wright, C. V. E. XBMPRII, a novel *Xenopus* type II receptor mediating BMP signalling in embryonic tissues. *Development* **125**, 431–442 (1998).
40. Hata, A., Lagna, G., Massague, J. & Hemmati-Brivanlou, A. SMAD6 inhibits BMP/SMAD1 signaling by specifically competing with the SMAD4 tumor suppressor. *Genes Dev.* **12**, 186–197 (1998).
41. Onichtchouk, D. *et al.* Silencing of TGF- β signalling by the pseudoreceptor BAMBI. *Nature* **400**, 480–485 (1999).
42. Wagner, G. P. In *Advances in Artificial Life* (eds Moran, F., Moreno, A., Merelo, J. J. & Chacon, P.) 317–328 (Springer, Berlin, 1995).

Acknowledgements

We thank D. Duboule and W. Pyerin for critical reading of the manuscript.

Correspondence should be addressed to C. N.

The DNA sequence of human chromosome 22

I. Dunham, N. Shimizu, B. A. Roe, S. Chissole *et al.*[†]

[†]A full list of authors appears at the end of this paper

Knowledge of the complete genomic DNA sequence of an organism allows a systematic approach to defining its genetic components. The genomic sequence provides access to the complete structures of all genes, including those without known function, their control elements, and, by inference, the proteins they encode, as well as all other biologically important sequences. Furthermore, the sequence is a rich and permanent source of information for the design of further biological studies of the organism and for the study of evolution through cross-species sequence comparison. The power of this approach has been amply demonstrated by the determination of the sequences of a number of microbial and model organisms. The next step is to obtain the complete sequence of the entire human genome. Here we report the sequence of the euchromatic part of human chromosome 22. The sequence obtained consists of 12 contiguous segments spanning 33.4 megabases, contains at least 545 genes and 134 pseudogenes, and provides the first view of the complex chromosomal landscapes that will be found in the rest of the genome.

Two alternative approaches have been proposed to determine the human genome sequence. In the clone by clone approach, a map of the genome is constructed using clones of a suitable size (for example, 100–200 kilobases (kb)), and then the sequence is determined for each of a representative set of clones that completely covers the map¹. Alternatively, a whole genome shotgun² requires the sequencing of unmapped genomic clones, typically in a size range of 2–10 kb, followed by a monolithic assembly to produce the entire sequence. Although the merits of these two strategies continue to be debated³, the public domain human genome sequencing project is following the clone by clone approach⁴ because it is modular, allows efficient organization of distributed resources and sequencing capacities, avoids problems arising from distant repeats and results in early completion of significant units of the genome. Here we report the first sequencing landmark of the human genome project, the operationally complete sequence of the euchromatic portion of a human chromosome.

Chromosome 22 is the second smallest of the human autosomes, comprising 1.6–1.8% of the genomic DNA⁵. It is one of five human acrocentric chromosomes, each of which shares substantial sequence similarity in the short arm, which encodes the tandemly repeated ribosomal RNA genes and a series of other tandem repeat sequence arrays. There is no evidence to indicate the presence of any protein coding genes on the short arm of chromosome 22 (22p). In contrast, direct⁶ and indirect^{7,8} mapping methods suggest that the long arm of the chromosome (22q) is rich in genes compared with other chromosomes. The relatively small size and the existence of a high-resolution framework map of the chromosome⁹ suggested to us that sequencing human chromosome 22 would provide an excellent opportunity to show the feasibility of completing the sequence of a substantial unit of the human genome. In addition, alteration of gene dosage on part of 22q is responsible for the aetiology of a number of human congenital anomaly disorders including cat eye syndrome (CES, Mendelian Inheritance in Man (MIM) 115470, <http://www.ncbi.nlm.nih.gov/omim/>) and velocardiofacial/DiGeorge syndrome (VCFS, MIM 192430; DGS, MIM 188400). Other regions associated with human disease are the schizophrenia susceptibility locus^{10,11}, and the sequences involved in spinocerebellar ataxia 10 (SCA10)¹². Making the sequence of human chromosome 22 freely available to the community early in the data collection phase has benefited studies of disease-related and other genes associated with this human chromosome^{13–19}.

Genomic sequencing

To identify genomic clones as the substrate for sequencing chromo-

some 22, extensive clone maps of the chromosome were constructed using cosmids, fosmids, bacterial artificial chromosomes (BACs) and P1-derived artificial chromosomes (PACs). Clones representing parts of chromosome 22 were identified by screening BAC and PAC libraries representing more than 20 genome equivalents using sequence tagged site (STS) markers known to be derived from the chromosome, or by using cosmid and fosmid libraries derived from flow-sorted DNA from chromosome 22. Overlapping clone contigs were assembled on the basis of restriction enzyme fingerprints and STS-content data, and ordered relative to each other using the established framework map of the chromosome⁹. The resulting nascent contigs were extended and joined by iterative cycles of chromosome walking using sequences from the end of each contig. In two places, yeast artificial chromosome (YAC) clones were used to join or extend contigs (AL049708, AL049760). The sequence-ready map covers 22q in 11 clone contigs with 10 gaps and stretches from sequences containing known chromosome 22 centromeric tandem repeats to the 22q telomere²⁰.

In the final sequence, one additional gap that was intractable to sequencing is found 234 kb from the centromere (see below). The gaps between the clone contigs are located at the two ends of the map, in the 4.3 Mb adjacent to the centromere and in 7.3 Mb at the telomeric end. These regions are separated by a central contig of 23 Mb. We have concluded that the gaps contain sequences that are unclonable with the available host-vector systems, as we were unable to detect clones containing the sequences in these gaps by screening more than 20 genome equivalents of bacterial clones using sequences adjacent to the contig ends.

The size of the seven gaps in the telomeric region has been estimated by DNA fibre fluorescence *in situ* hybridization (FISH). No gap in this region is judged to be larger than ~150 kb. For three of these gaps, a number of BAC and PAC clones that contain STSs on either side of the gap were shown to be deleted for at least a minimal core region by DNA fibre FISH. As these clones come from multiple donor DNA sources, these results are unlikely to be due to deletion in the DNA used to make the libraries. Furthermore, the same result was observed for the gap at 32,600 kb from the centromeric end of the sequence, when the DNA fibre FISH experiments were performed on DNA from two different lymphoblastoid cell lines. One possible explanation for this observation is that DNA fragments containing the gap sequences are initially cloned in the BAC library but clones that delete these sequences have a significant selective advantage as the library is propagated. As the observed size range of the cloned inserts in the BAC libraries ranges from 100 kb to more than 230 kb (<http://bacpac.med.buffalo.edu/>), such deletion events

are not distinguishable on the basis of size from undeleted BACs. Additional analysis of the distribution of BAC end sequences from dbGSS (<http://www.ncbi.nlm.nih.gov/dbGSS/index.html>) suggests that the BAC coverage is sparser closer to the gaps and that this analysis did not identify any BACs spanning the gaps. The three remaining clone-map gaps in the proximal region of the long arm are in regions that may contain segments of previously characterized low-copy repeats²¹. These gaps could not be sized by DNA fibre FISH because of the extensive intra- and interchromosomal repeat sequences (see below) but were amenable to long-range restriction mapping. The gap between AP000529 and AP000530 was estimated to be shorter than 150 kb by comparison with a previously established long-range restriction map²². The gap closest to the centromere, which is less than 2 kb in size, could not be sequenced despite BAC clone coverage as it was unrepresented in plasmid or M13 libraries, and was intractable to all sequencing strategies applied. Detailed descriptions of several of the clone contigs have been published^{21,23,24} or will be published elsewhere.

Each sequencing group took responsibility for completion of adjacent areas of the sequence as illustrated in Fig. 1. A set of minimally overlapping clones (the 'tile path') was chosen from the physical map and sequenced using a combination of a random shotgun assembly, followed by directed sequencing to close gaps and resolve ambiguities ('finishing'). The major problems encountered during completion of the sequence in the directed sequencing phase were CpG islands, tandem repeats and apparent cloning biases. Directed sequencing using oligonucleotide primers, very short insert plasmid libraries, or identification of bridging clones by screening high complexity plasmid or M13 libraries solved these problems.

The completed sequence covers 33.4 Mb of 22q with 11 gaps and has been estimated to be accurate to less than 1 error in 50,000 bases, by internal and external checking exercises²⁵. The order and size of each of the contiguous pieces of sequence is detailed in Table 1. The largest contiguous segment stretches over 23 Mb. From our gap-size estimates, we calculate that we have completed 33,464 kb of a total region spanning 34,491 kb and that therefore the sequence is complete to 97% coverage of 22q. The complete sequence and analysis is available on the internet (<http://www.sanger.ac.uk/HGP/Chr22> and <http://www.genome.ou.edu/Chr22.html>).

Sequence analysis and gene content

Analysis of the genomic sequence of the model organisms has made extensive use of predictive computational analysis to identify genes^{26–28}. In human DNA, identification of genes by these methods is more difficult because of extensive splicing, lower density of exons and the high proportion of interspersed repetitive sequences. The accuracy of *ab initio* gene prediction on vertebrate genomic sequence has been difficult to determine because of the lack of sequence that has been completely annotated by experiment. To determine the degree of overprediction made by such algorithms, all genes within a region need to be experimentally identified and annotated, however it is virtually impossible to know when this job is complete. A 1.4-Mb region of human genomic sequence around the BRCA2 locus has been subjected to extensive experimental investigation, and it is believed that the 170 exons identified is close to the total number expressed in the region.

The most recent calibration of *ab initio* methods against this region (R.B.S.K. and T.H., manuscript in preparation) shows that with the best methods^{29,30} more than 30% of exon predictions do not overlap any experimental exons, in other words, they are over-predictions. Furthermore, having now applied this analysis to larger amounts of data (more than 15 Mb from the Sanger Annotated Genome Sequence Repository which can be obtained as part of the Genesafe collection (<http://www.hgmp.mrc.ac.uk/Genesafe/>)), it is confirmed that prediction accuracy also varies considerably between different regions of sequence. It was hoped that these calibration efforts would lead to rules for reliable gene prediction

based on *ab initio* methods alone, perhaps on the basis of combining several different methods, GC content and so on. However, so far this has not been possible. The same analysis also shows that although ~95% of genes are at least partially predicted by *ab initio* methods, few gene structures are completely correct (none in BRCA2) and more than 20% of experimental exons are not predicted at all. The comparison of *ab initio* predictions and the annotated gene structures (see below) in the chromosome 22 sequence is consistent with this, with 94% of annotated genes at least partially detected by a Genscan gene prediction, but only 20% of annotated genes having all exons predicted exactly. Sixteen per cent of all the exons in annotated genes were not predicted at all, although this is only 10% for internal exons (that is, not 5' and 3' ends). As a result, we do not consider that *ab initio* gene prediction software can currently be used directly to reliably annotate genes in human sequence, although it is useful when combined with other evidence (see below), for example, to define splice-site boundaries, and as a starting point for experimental studies.

Fortunately, a vast resource of experimental data on human genes in the form of complementary DNA and protein sequences and expressed sequence tags (ESTs) is available which can be used to identify genes within genomic DNA. Furthermore about 60% of human genes have distinctive CpG island sequences at their 5' ends³¹ which can also be used to identify potential genes. Thus, the approach we have taken to annotating genes in the chromosome 22 sequence relies on a combination of similarity searches against all available DNA and protein databases, as well as a series of *ab initio* predictions. Upon completion of the sequence of each clone in the tile path, the sequence was subjected to extensive computational analysis using a suite of similarity searches and prediction tools. Briefly, the sequences were analysed for repetitive sequence content, and the repeats were masked using RepeatMasker (<http://ftp.genome.washington.edu/RM/RepeatMasker.html>). Masked sequence was compared to public domain DNA and protein databases by similarity searches using the blast family of programs³². Unmasked sequence was analysed for C + G content and used to predict the presence of CpG islands, tandem repeat sequences, tRNA genes and exons. The completed analysis was assembled into contigs and visualized using implementations of ACEDB (<http://www.sanger.ac.uk/Software/Acedb/>). In addition, the contiguous masked sequence was analysed using gene prediction software^{29,30}.

Figure 1 The sequence of human chromosome 22. Coloured boxes depict the annotated features of the sequence of human chromosome 22, with the centromere to the left and the telomere to the right. Coordinates are in kilobases. Vertical yellow blocks indicate the positions of the gaps in the sequence and are proportional in size to the estimated size of each gap. From bottom to top the following features are displayed: positions of interspersed repetitive sequences including tandem repeats categorized by nucleotide repeat unit length (at this resolution *Alu* repeats are not visibly separated in some regions); the positions of the microsatellite markers in the genetic map of Dib *et al.*³⁶; the tiling path of genomic clones used to determine the sequence labelled by their Genbank/EMBL/DBJ accession number and coloured according to the source of the sequence; and the annotated gene, pseudogene and CpG island content of the sequence. Transcripts and pseudogenes oriented 5' to 3' on the DNA strand from centromere to telomere are designated '+', those on the opposite strand '-'. In the transcript rows, the annotated genes are subdivided by colour according to the criteria in the text. Annotated genes with approved gene symbols from the HUGO nomenclature committee are labelled. For details of all the genes with their positions in the reference sequence, see Supplementary Information, Table 1. In the case of the immunoglobulin variable region, the entire locus has been drawn as a single block; in reality, this is a complex of variable chain genes (see ref. 27). At the top is a graphical plot of the repeat density for the common interspersed repeats *Alu* and *Line1*, and the C + G base frequency across the sequence. Each is calculated as a percentage of the sequence using a sliding 100-kb window moved in 50-kb iterations. Since the production of Fig. 1, six accession codes have been updated. The new codes are AL050347 (for Z73987), AL096754 (for Z68686), AL049749 (for Z82197), Z75892 (for Z75891), AL078611 (for Z79997) and AL023733 (for AL023593).

Gene features were identified by a combination of human inspection and software procedures. Figure 1 shows the 679 gene sequences annotated across 22q. They were grouped according to the evidence that was used to identify them as follows: genes identical to known human gene or protein sequences, referred to as 'known genes' (247); genes homologous, or containing a region of similarity, to gene or protein sequences from human or other species, referred to as 'related genes' (150); sequences homologous to only ESTs, referred to as 'predicted genes' (148); and sequences homologous to a known gene or protein, but with a disrupted open reading frame, referred to 'pseudogenes' (134). (See Supplementary Information, Table 1, for details of these genes.) The *ab initio* gene prediction program, Genscan, predicted 817 genes (6,684 exons) in the contiguous sequence, of which 325 do not form part of the annotated genes categorized above. Given the calibration of *ab initio* prediction methods discussed above, we estimate that of the order of 100 of these will represent parts of 'real' genes for which there is currently no supporting evidence in any sequence database, and that the remainder are likely to be false positives.

The total length of the sequence occupied by the annotated genes, including their introns, is 13.0 Mb (39% of the total sequence). Of this, only 204 kb contain pseudogenes. About 3% of the total sequence is occupied by the exons of these annotated genes. This contrasts sharply with the 41.9% of the sequence that represents tandem and interspersed repeat sequences. There is no significant bias towards genes encoded on one strand at the 5% level ($\chi^2 = 3.83$).

A striking feature of the genes detected is their variety in terms of both identity and structure. There are several gene families that appear to have arisen by tandem duplication. The immunoglobulin λ locus is a well-known example, but there also are other immunoglobulin-related genes on the chromosome outside the immunoglobulin λ region. These include the three genes of the immunoglobulin λ -like (IGLL) family plus a fourth possible member of the family (AC007050.7). There are five clustered immunoglobulin κ variable region pseudogenes in AC006548, and an immunoglobulin variable-related sequence (VpreB3) in AP000348. Much further away from the λ genes is a variable region pseudogene, ~123 kb telomeric of IGLL3 in sequence AL008721 (coordinates ~9,420–9,530 kb from the centromeric

end of the sequence), and a cluster of two λ constant region pseudogenes and a variable region pseudogene in sequences AL008723/AL021937 (coordinates ~16,060–16,390 kb from the centromeric end).

Human chromosome 22 also contains other duplicated gene families that encode glutathione S-transferases, Ret-finger-like proteins¹⁹, phorbolins or APOBECs, apolipoproteins and β -crystallins. In addition, there are families of genes that are interspersed among other genes and distributed over large chromosomal regions. The γ -glutamyl transferase genes represent a family that appears to have been duplicated in tandem along with other gene families, for instance the BCR-like genes, that span the 22q11 region and together form the well-known LCR22 (low-copy repeat 22) repeats (see below).

The size of individual genes encoded on this chromosome varies over a wide range. The analysis is incomplete as not all 5' ends have been defined. However, the smallest complete genes are only of the order of 1 kb in length (for example, HMG1L10 is 1.13 kb), whereas the largest single gene (LARGE¹⁵) stretches over 583 kb. The mean genomic size of the genes is 19.2 kb (median 3.7 kb). Some complete gene structures appear to contain only single exons, whereas the largest number of exons in a gene (PIK4CA) is 54. The mean exon number is 5.4 (median 3). The mean exon size is 266 bp (median 135 bp). The smallest complete exon we have identified is 8 bp in the PITPNB gene. The largest single exon is 7.6 kb in the PKDREJ, which is an intron-less gene with a 6.7-kb open reading frame. In addition, two genes occur within the introns of other expressed genes. The 61-kb TIMP3 gene, which is involved in Sorsby fundus macular degeneration, lies within a 268-kb intron of the large SYN3 gene, and the 8.5 kb HCF2 gene lies within a 27.5-b intron of the PIK4CA gene. In each case, the genes within genes are oriented in the opposite transcriptional orientation to the outer gene. We also observe pseudogenes frequently lying within the introns of other functional genes.

Peptide sequences for the 482 annotated full-length and partial genes with an open reading frame of greater than or equal to 50 amino acids were analysed against the protein family (PFAM)³³, Prosite³⁴ and SWISS-PROT³⁵ databases. These data were processed and displayed in an implementation of ACEDB. Overall, 240 (50%) predicted proteins had matching domains in the PFAM database encompassing a total of 164 different PFAM domains. Of the residues making up these 482 proteins, 25% were part of a PFAM domain. This compares with PFAM's residue coverage of SWISS-PROT/TrEMBL, which is more than 45% and indicates that the human genome is enriched in new protein sequences. Sixty-two PFAM domains were found to match more than one protein, including ten predicted proteins containing the eukaryotic protein kinase domain (PF00069), nine matching the Src homology domain 3 (PF00018) and eight matching the RhoGAP domain (PF00620). Fourteen predicted proteins contain zinc-finger domains (See

Table 1 Sequence contigs on chromosome 22

Contig*	Size (kb)	
AP000522-AP000529	234	1.9
gap		
AP000530-AP000542	406	~150
gap		
AP000543-AC006285	1,394	~150
gap		
AC008101-AC007663	1,790	~100
gap		
AC007731-AL049708	23,006	~50
gap		
AL118498-AL022339	767	~50-100
gap†		
Z85994-AL049811	1,528	~150
gap		
AL049853-AL096853	2,485	~50
gap†		
AL096843-AL078607	190	~100
gap†		
AL078613-AL117328	993	~100
gap		
AL080240-AL022328	291	~100
gap†		
AL096767-AC002055	380	
Total sequence length	33,464	
Total length of 22q	34,491	

* Contigs are indicated by the first and last sequence in the orientation centromere to telomere, and are named by their Genbank/EMBL/DBJ accession numbers.

† These gaps are spanned by BAC and/or PAC clones with deletions.

‡ This gap shows a complex duplication of AL096853 in DNA fibre FISH.

Table 2 The interspersed repeat content of human chromosome 22

Repeat type	Total number	Coverage (bp)	Coverage (%)
Alu	20,188	5,621,998	16.80
HERV	255	160,697	0.48
Line1	8,043	3,256,913	9.73
Line2	6,381	1,273,571	3.81
LTR	848	256,412	0.77
MER	3,757	763,390	2.28
MIR	8,426	1,063,419	3.18
MLT	2,483	605,813	1.81
THE	304	93,159	0.28
Other	2,313	625,562	1.87
Dinucleotide	1,775	133,765	0.40
Trinucleotide	166	18,410	0.06
Quadranucleotide	404	47,691	0.14
Pentanucleotide	16	1,612	0.0048
Other tandem	305	102,245	0.31
Total	55,664	14,024,657	41.91

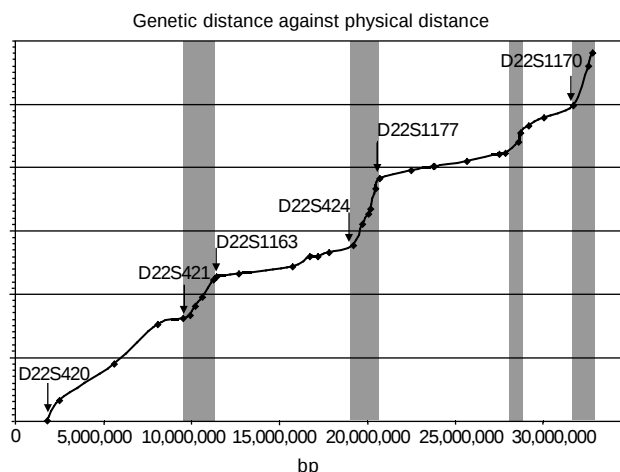


Figure 2 The relationship between physical and genetic distance. The sex-averaged genetic distances of Dib *et al.*³⁶ were obtained from <ftp://ftp.genethon.fr/pub/Gmap/Nature-1995/> and the cumulative intermarker distances for unambiguously ordered markers (in cM) were plotted against the positions of the microsatellite markers in the genomic sequence. It should be stressed that the y axis does not represent the true genetic distance between distant markers but the sum of the local intermarker distances. The positions of selected genetic markers are labelled. Grey regions are indicative of areas of relatively increased recombination per unit physical distance.

Supplementary Information, Table 2, for details of the PFAM domains identified in the predicted proteins).

Nineteen per cent of the coding sequences identified were designated as pseudogenes because they had significant similarity to known genes or proteins but had disrupted protein coding reading frames. Because 82% of the pseudogenes contained single blocks of homology and lacked the characteristic intron–exon structure of the putative parent gene, they probably are processed pseudogenes. Of the remaining spliced pseudogenes, most represent segments of duplicated gene families such as the immunoglobulin κ variable genes, the β -crystallins, CYP2D7 and CYP2D8, and the GGT and BCR genes. The pseudogenes are distributed over the entire sequence, interspersed with and sometimes occurring within the introns of annotated expressed genes. However, there also is a dense cluster of 26 pseudogenes in the 1.5-Mb region immediately adjacent to the centromere; the significance of this cluster is currently unclear.

Given that the sequence of 33.4 Mb of chromosome 22q represents 1.1% of the genome and encodes 679 genes, then, if the distribution of genes on the other chromosomes is similar, the minimum number of genes in the entire human genome would be at least 61,000. Previous work has suggested that chromosome 22 is gene rich⁶ by a factor of 1.38 (<http://www.ncbi.nlm.nih.gov/genemap/page.cgi?F=GeneDistrib.html>), which would reduce this estimate to ~45,000 genes. It is important, however, to recognize that the analysis described here only provides a minimum estimate for the gene content of chromosome 22q, and that further studies will probably reveal additional coding sequences that could not be identified with the current approaches.

Two lines of evidence point to the existence of additional genes that are not detected in this analysis. First, the 553 predicted CpG islands, which typically lie at the true 5' ends of about 60% of human genes³¹, are in excess of 60% of the number of genes identified (60% = 327, excluding pseudogenes); 282 of the genes identified have CpG islands at or close to the 5' end (within 5-kb upstream of the first exon, or 1-kb downstream). Thus, there could be up to 271 additional genes associated with CpG islands undetected in the sequence. Second, there are 325 putative genes predicted by the *ab initio* gene prediction program, Genscan, that are not in regions already containing annotated transcripts. We

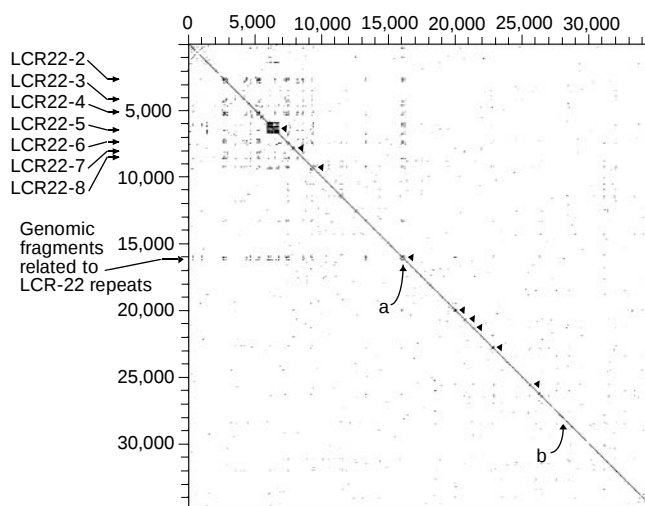


Figure 3 Intrachromosomal repeats on human chromosome 22. High- and medium-copy repeats and low complexity sequence were masked using RepeatMasker and Dust, and masked sequences were compared using Blastn. The results were filtered to identify regions of more than 50% identity to the query sequence, and were plotted in a 2D matrix with a line proportional to the size of the region of identity. Localized gene family repeats are indicated by arrowheads along the diagonal. From the top, these are the immunoglobulin λ locus, the glutathione S-transferase genes, the β -crystallin genes, the Ret-finger-protein-like genes, the apolipoprotein genes, the colony-stimulating factor receptor (CSF2RB) inverted partial duplication, the lectins LGALS1 and LGALS2, the APOBEC genes and the CYP2D genes. Two 60-kb regions of more than 90% homology are labelled 'a' (AL008723/AL021937) and 'b' (AL031595/AL022339). Seven low-copy repeat regions (LCR22) and a region containing related genomic fragments are indicated at the left margin.

estimate (see above) that roughly 100 of these will represent parts of real genes. Identifying additional genes will require further computational and experimental studies. These studies are continuing and entail testing candidate sequences for possible messenger RNA expression, implementing new gene prediction software able to detect the regions around or near CpG islands that currently have no identified transcript, and further analysis of sequences that are conserved between human and mouse. Furthermore, full-length cDNA sequences that accumulate in the sequence databases of human and other species will be used to refine the gene structures.

The long-range chromosome landscape

Critical to the utility of the genomic sequence to genetic studies is the integration of established genetic maps. The positions of the commonly used microsatellite markers from the Genethon genetic map³⁶ are given in Fig. 1. The correlation of the order of markers between the genetic map and the sequence is good, within the limitations of genetic mapping. Only a single marker (D22S1175) is discrepant between the two data sets, and this lies in a sequence that is repeated twice on the chromosome (AL021937, see below). In the telomeric region, four of the Genethon markers must lie in our sequence gaps, and we were unable to identify clones from all libraries tested for these. Comparison of genetic distance against physical distance for all the microsatellites whose order is maintained between the datasets shows a mean value of 1.87 cM Mb⁻¹. However, the relationship between genetic and physical distance across the chromosome partitions into two types of region, areas of high and low recombination (Fig. 2). The areas of high recombination may represent recombinational hot spots, although we have not yet been able to identify any specific sequence characteristics common to these areas.

The mean G + C content of the sequence is 47.8%. This is significantly higher than the G + C content calculated for the

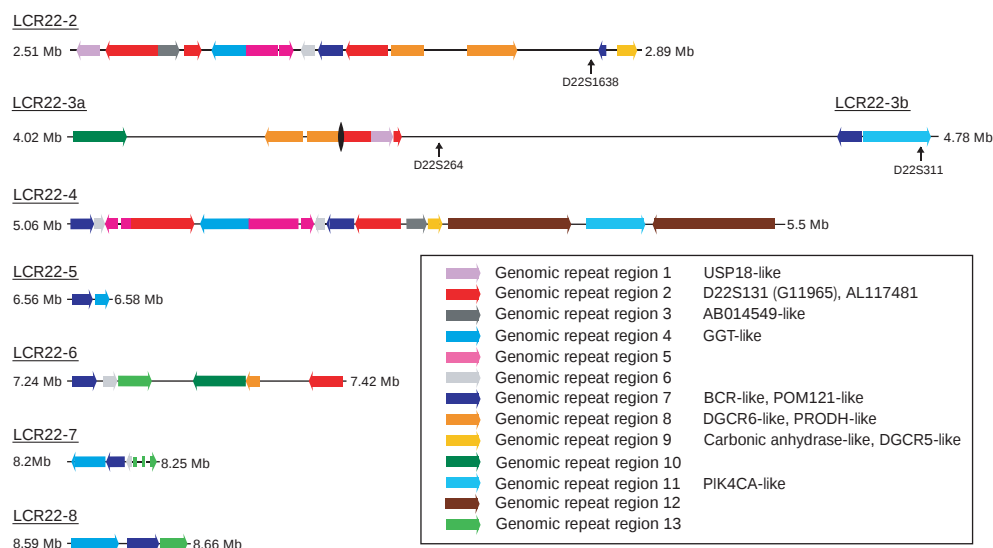


Figure 4 Sequence composition of the LCR22 repeats. Illustration of the sequence composition of seven LCR22 repeats. The span of each LCR22 region is shown in megabases from the centromere. Coloured arrows indicate the extent of one of the

thirteen genomic repeat regions and the orientation of the repeat. The known gene and marker content of these genomic repeat regions is indicated in the key. The black oval indicates the position of the gap in the sequence in LCR22-3.

sum of all human genomic sequence determined so far (42%). Although this result was expected from previous indirect measurements of the G + C content of chromosome 22^{7,8,37}, the distribution is not uniform, but regionally segmented as illustrated in Fig. 1. There are clear fluctuations in the base content, resulting in areas that are relatively G + C rich and others that are relatively G + C poor. On chromosome 22 these regions stretch over several megabases. For example, the 2 Mb of sequence closest to the centromeric end of the sequence is relatively G + C poor, with the G + C content dropping below 40%. Similarly, the area between 16,000 and 18,800 kb from the centromeric end of the sequence is consistently below 45% G + C. The G + C rich regions often reach more than 55% G + C (for example, at 20,100–23,400 kb from the centromeric end of the sequence). This fluctuation appears to be consistent with previous observations that vertebrate genomes are segmented into ‘isochores’ of distinct G + C content³⁸ and is similar to the structure seen in the human major histocompatibility complex (MHC) sequence³⁹. Isochores correlate with both genes and chromosome structure. The G + C rich isochores are rich in genes and *Alu* repeats, and are located in the G + C rich chromosomal R-bands, whereas the G + C poor isochores are relatively depleted in genes and *Alu* repeats, and are located in the G-bands^{8,37,40}. The G + C poor regions of chromosome 22 are depleted in genes and relatively poor in *Alu* sequences. For example, the region between 16,000 and 18,800 kb from the centromeric end contains just three genes, two of which are greater than 400 kb in length. The G + C poor regions also are depleted in CpG islands, which are clustered in the gene-rich, G + C rich regions. Although it is tempting to correlate the sequence features that we see with the chromosome banding patterns, we believe that high-resolution mapping of the chromosome band boundaries will be required to assign definitively these to genomic sequence.

Over 41.9% of the chromosome 22 sequence comprises interspersed and tandem repeat family sequences (Table 2). The density of repeats across the sequence is plotted in Fig. 1. There is variation in the density of *Alu* repeats and some of the regions with low *Alu* density correlate with the G + C poor regions, for example, in the region 16,000–18,800 kb from the centromeric end, and these data support the relationship of isochores with *Alu* distribution. However, in other areas the relationship is less clear. We provide a World-Wide Web interface to the long-range analyses presented here and to further analysis of the many other repeat types and features of the

sequence at <http://www.sanger.ac.uk/cgi-bin/cwa/22cwa.pl>. The 1-Mb region closest to the centromere contains several interesting repeat sequence features that may be typical of other pericentromeric regions. In addition to the density of pseudogenes described above, there is a large 120-kb block of tandemly repeated satellite sequence (D22Z3) centred 500 kb from the centromeric sequence start (not shown in Fig. 1, but evident from the absence of *Alu* and LINE1 sequences at this point). There is also a cluster of satellite II repeats 80-kb telomeric of the D22Z3 sequences. Isolated alphoid satellite repeats are found closer to the centromeric end of the sequence. Furthermore, this pericentromeric 1 Mb closest to the centromere contains many sequences that are shared with a number of different chromosomes, particularly chromosomes 2 and 14. During map construction, 33 out of 37 STSs designed from sequence that was free of high-copy repeats amplified from more than one chromosome in somatic cell hybrid panel analysis.

Low-copy repeats on chromosome 22

To detect intra- and interchromosomal repeats, we compared the entire sequence of chromosome 22 to itself, and also to all other existing human genomic DNA sequence using Blastn³² after masking high and medium frequency repeats. The results of the intra-chromosomal sequence analysis were plotted as a dot matrix (Fig. 3) and reveal a series of interesting features. Locally duplicated gene families lie close to the diagonal axis of the plot. The most striking is the immunoglobulin λ locus that comprises a cluster of 36 potentially functional V- λ gene segments, 56 V- λ pseudogenes, and 27 partial V- λ pseudogenes (‘relics’), together with 7 each of the J and C λ segments²⁴. Other duplicated gene families that are visible from the dot matrix plot include the clustered genes for glutathione S-transferases, β -crystallins, apolipoproteins, phorbolins or APO-BECs, the lectins LGALS1 and LGALS2 and the CYP2Ds. A partial inverted duplication of CSF2RB is also observed.

Much more striking are the long-range duplications, which are visible away from the diagonal axis. For example, a 60-kb segment of more than 90% similarity is seen between sequences AL008723/AL021937 (at ~16,060–16,390 kb from the centromeric end) and AL031595/AL022339 (at ~27,970–28,110 kb from the centromeric end) separated by almost 12 Mb. The 22q11 region is particularly rich in repeated clusters⁴¹. Previous work described a low-copy repeat family in 22q11 that might mediate recombination events leading to the chromosomal rearrangements seen in cat eye,

velocardiofacial and DiGeorge syndromes^{21,42}. The availability of the entire DNA sequence allows detailed dissection of the molecular structure of these low-copy repeats (LCR22s). Edelman *et al.* described eight LCR22 regions^{21,42}. We were unable to find the LCR22 repeat closest to the centromere, but it may lie in the gap at 700 kb from the centromeric end of the sequence. The other LCR22 regions are distributed over 6.5 Mb of 22q11. Analysis of the sequence shows that each LCR22 contains a set of genes or pseudogenes (Fig. 4). For example, five of the LCR22s contain copies of the γ -glutamyl transferase genes and γ -glutamy-transferase-related genes. There is also evidence that a more distant sequence at ~16,000 kb from the centromeric start of the genomic sequence shares certain sequences with the LCR22 repeats. This similarity involves related genomic fragments including parts of the Ret-finger-protein-like genes, and the IGLC and IGLV genes.

Regions of conserved synteny with the mouse

The genomic organization of different mammalian species is well known to be conserved⁴³. Comparison of genetic and physical maps across species can aid in predicting gene locations in other species, identifying candidate disease genes¹³, and revealing various other features relevant to the study of genome organization and evolution. For all the cross-species relationships, that between man and mouse has been most studied. We have examined the relationship of the human chromosome 22 genes to their mouse orthologues.

Of the 160 genes we identified in the human chromosome 22 sequence that have orthologues in mouse, 113 of the murine orthologues have known mouse chromosomal locations (data available at <http://www.sanger.ac.uk/HGP/Chr22/Mouse/>). Examination of these mouse chromosomal locations mapped onto the human chromosome 22 sequence confirms the conserved linkage groups corresponding to human chromosome 22 on mouse chro-

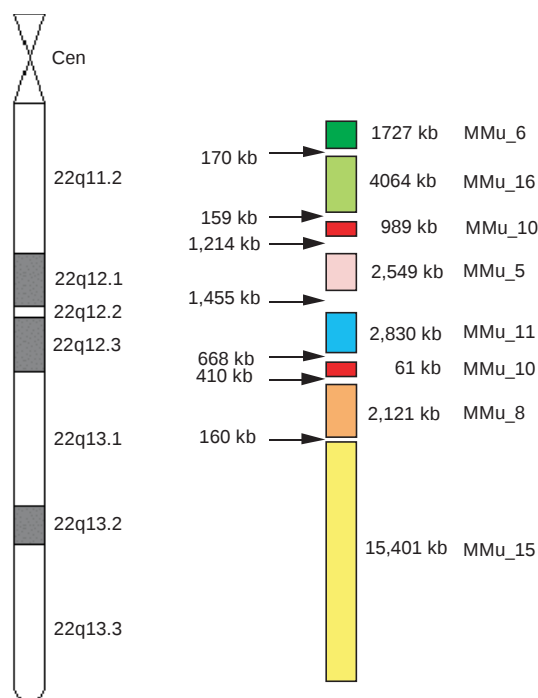


Figure 5 Regions of conserved synteny between human chromosome 22 and the mouse genome. Regions of mouse chromosomes with conserved synteny to human chromosome 22 are shown as adjacent coloured blocks, determined by the mouse map position of mouse orthologues to human chromosome 22 genes. The size of human chromosome 22 corresponding to each mouse chromosomal region is indicated in kb, as well as the size of the gap between the last orthologue in each conserved block. These data are available at <http://www.sanger.ac.uk/Chr22/Mouse>.

mosomes 6, 16, 10, 5, 11, 8 and 15^{18,44–46} (Fig. 5). Furthermore, these studies allow placement of the sites of evolutionary rearrangements that have disrupted the conservation of synteny more accurately at the DNA sequence scale. For example, the breakdown of synteny between the equivalents of the human HMOX1 and MB genes, which are separated by less than 160 kb that also contains a conspicuous 41-copy 18-nucleotide tandem repeat. A clear prediction from these data is that, for the most part, the unmapped murine orthologues of the human genes lie within these established linkage groups, along with the orthologues of the human genes that currently lack mouse counterparts. Exploitation of the chromosome 22 sequence may hasten the determination of the mouse genomic sequence in these regions.

Conclusions

We have shown that the clone by clone strategy is capable of generating long-range continuity sufficient to establish the operationally complete genomic sequence of a chromosome. In doing so, we have generated the largest contiguous segment of DNA sequence to our knowledge to date. The analysis of the sequence gives a foretaste of the information that will be revealed from the remaining chromosomes.

We were unable to obtain sequence over 11 small gaps using the available cloning systems. It may be possible that additional approaches such as using combinations of cloning systems with small insert sizes and low-copy number could reduce the size of these gaps. Direct cloning of restriction fragments that cross these gaps into small insert plasmid or M13 libraries, or direct sequencing approaches might eventually provide access to all the sequence in the gaps. However, closing these gaps is certain to require considerable time and effort, and might be considered as a specialist activity outside the core genome-sequencing efforts. It also is probable that the sequence features responsible for several of these gaps are unlikely to be specific to chromosome 22. In the best case, similar unclonable sequences might be restricted to the centromeric and telomeric regions of the other chromosomes and areas with large tandem repeats, and it will be possible to obtain large contiguous segments for the bulk of the euchromatic genome.

Over the course of the project, the emerging sequence of chromosome 22 has been made available in advance of its final completion through the internet sites of the consortium groups and the public sequence databases⁴⁷. The benefits of this policy can be seen in both the regular requests received from investigators for materials and information that arise as the result of sequence homology searches, and the publications that have used the data^{14–19}. The genome project will continue to pursue this data release policy as we move closer to the anticipated completed sequence of humans, mice and other complex genomes^{47,48}.

Methods

The methods for construction of clone maps have been previously described^{24,49,50} and can also be found at <http://www.sanger.ac.uk/HGP/methods/>. Details of sequencing methods and software are available at <http://www.sanger.ac.uk/HGP/methods/>, <http://www.genome.ou.edu/proto.html>, http://www-alis.tokyo.jst.go.jp/HGS/team_KU/team.html and in the literature^{1,24}.

Received 5 November; accepted 11 November 1999.

1. The Sanger Centre & The Genome Sequencing Centre. Toward a complete human genome sequence. *Genome Res.* **8**, 1097–1108 (1998).
2. Weber, J. L. & Myers, E. W. Human whole-genome shotgun sequencing. *Genome Res.* **7**, 401–409 (1997).
3. Green, P. Against a whole-genome shotgun. *Genome Res.* **7**, 410–417 (1997).
4. Collins, F. S. *et al.* New goals for the U.S. human genome project: 1998–2003. *Science* **282**, 682–689 (1998).
5. Morton, N. E. Parameters of the human genome. *Proc. Natl Acad. Sci. USA* **88**, 7474–7476 (1991).
6. Deloukas, P. *et al.* A physical map of 30,000 human genes. *Science* **282**, 744–746 (1998).
7. Craig, J. M. & Bickmore, W. A. The distribution of CpG islands in mammalian chromosomes. *Nature Genet.* **7**, 376–382 (1994).
8. Saccone, S., Caccio, S., Kusuda, J., Andreozzi, L. & Bernardi, G. Identification of the gene-rich bands in human chromosomes. *Gene* **174**, 85–94 (1996).

9. Collins, J. E. *et al.* A high-density YAC contig map of human chromosome 22. *Nature* **377**, 367–379 (1995).
10. Pulver, A. E. *et al.* Psychotic illness in patients diagnosed with velo-cardio-facial syndrome and their relatives. *J. Nerv. Ment. Dis.* **182**, 476–478 (1994).
11. Gill, M. *et al.* A combined analysis of D22S278 marker alleles in affected sib-pairs: support for a susceptibility locus for schizophrenia at chromosome 22q12. Schizophrenia Collaborative Linkage Group (Chromosome 22). *Am. J. Med. Genet.* **67**, 40–45 (1996).
12. Zu, L., Figueroa, K. P., Grewal, R. & Pulst, S. M. Mapping of a new autosomal dominant spinocerebellar ataxia to chromosome 22. *Am. J. Hum. Genet.* **64**, 594–599 (1999).
13. Southard-Smith, E. M. *et al.* Comparative analyses of the dominant megacolon-SOX10 genomic interval in mouse and human. *Mamm. Genome* **10**, 744–749 (1999).
14. Nishino, I., Spinazzola, A. & Hirano, M. Thymidine phosphorylase gene mutations in MNGIE, a human mitochondrial disorder. *Science* **283**, 689–692 (1999).
15. Peyrard, M. *et al.* The human LARGE gene from 22q12.3-q13.1 is a new, distinct member of the glycosyltransferase gene family. *Proc. natl Acad. Sci. USA* **96**, 598–603 (1999).
16. Kao, H. T. *et al.* A third member of the synapsin gene family. *Proc. Natl Acad. Sci. USA* **95**, 4667–4672 (1998).
17. Mittman, S., Guo, J., Emerick, M. C. & Agnew, W. S. Structure and alternative splicing of the gene encoding alpha1I, a human brain T calcium channel alpha1 subunit. *Neurosci. Lett.* **269**, 121–124 (1999).
18. Seroussi, E. *et al.* TOM1 genes map to human chromosome 22q13.1 and mouse chromosome 8C1 and encode proteins similar to the endosomal proteins HGS and STAM. *Genomics* **57**, 380–388 (1999).
19. Seroussi, E. *et al.* Duplications on human chromosome 22 reveal a novel ret finger protein-like gene family with sense and endogenous antisense transcripts. *Genome Res.* **9**, 803–814 (1999).
20. Ning, Y., Rosenberg, M., Biesecker, L. G. & Ledbetter, D. H. Isolation of the human chromosome 22q telomere and its application to detection of cryptic chromosomal abnormalities. *Hum. Genet.* **97**, 765–769 (1996).
21. Edelmann, L., Pandita, R. K. & Morrow, B. E. Low-copy repeats mediate the common 3-Mb deletion in patients with velo-cardio-facial syndrome. *Am. J. Hum. Genet.* **64**, 1076–1086 (1999).
22. McDermid, H. E. *et al.* Long-range mapping and construction of a YAC contig within the cat eye syndrome critical region. *Genome Res.* **6**, 1149–1159 (1996).
23. Johnson, A. *et al.* A 1.5-Mb contig within the cat eye syndrome critical region at human chromosome 22q11.2. *Genomics* **57**, 306–309 (1999).
24. Kawasaki, K. *et al.* One-megabase sequence analysis of the human immunoglobulin lambda gene locus. *Genome Res.* **7**, 250–261 (1997).
25. Felsenfeld, A., Peterson, J., Schloss, J. & Guyer, M. Assessing the quality of the DNA sequence from the Human Genome Project. *Genome Res.* **9**, 1–4 (1999).
26. Mewes, H. W. *et al.* Overview of the yeast genome. *Nature* **387**, 7–65 (1997).
27. Blattner, F. R. *et al.* The complete genome sequence of *Escherichia coli* K-12. *Science* **277**, 1453–1474 (1997).
28. The C. elegans Sequencing Consortium. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* **282**, 2012–2018 (1998).
29. Solovyev, V. & Salamov, A. The Gene-Finder computer tools for analysis of human and model organisms genome sequences. *Ismb* **5**, 294–302 (1997).
30. Burge, C. & Karlin, S. Prediction of complete gene structures in human genomic DNA. *J. Mol. Biol.* **268**, 78–94 (1997).
31. Cross, S. H. & Bird, A. P. CpG islands and genes. *Curr. Opin. Genet. Dev.* **5**, 309–314 (1995).
32. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
33. Bateman, A. *et al.* Pfam 3.1: 1313 multiple alignments and profile HMMs match the majority of proteins. *Nucleic Acids Res.* **27**, 260–262 (1999).
34. Hofmann, K., Bucher, P., Falquet, L. & Bairoch, A. The PROSITE database, its status in 1999. *Nucleic Acids Res.* **27**, 215–219 (1999).
35. Bairoch, A. & Apweiler, R. The SWISS-PROT protein sequence data bank and its supplement TrEMBL in 1999. *Nucleic Acids Res.* **27**, 49–54 (1999).
36. Dib, C. *et al.* A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature* **380**, 152–154 (1996).
37. Holmquist, G. P. Chromosome bands, their chromatin flavors, and their functional features. *Am. J. Hum. Genet.* **51**, 17–37 (1992).
38. Bernardi, G. The isochore organization of the human genome. *Annu. Rev. Genet.* **23**, 637–661 (1989).
39. The MHC sequencing consortium. Complete sequence and gene map of a human major histocompatibility complex. *Nature* **401**, 921–923 (1999).
40. Bernardi, G. The isochore organization of the human genome. *Annu. Rev. Genet.* **23**, 637–661 (1989).
41. Collins, J. E., Mungall, A. J., Badcock, K. L., Fay, J. M. & Dunham, I. The organization of the gamma-glutamyl transferase genes and other low copy repeats in human chromosome 22q11. *Genome Res.* **7**, 522–531 (1997).
42. Edelman, L. *et al.* A common molecular basis for rearrangement disorders on chromosome 22q11. *Hum. Mol. Genet.* **8**, 1157–1167 (1999).
43. Eppig, J. T. & Nadeau, J. H. Comparative maps: the mammalian jigsaw puzzle. *Curr. Opin. Genet. Dev.* **5**, 709–716 (1995).
44. Bucan, M. *et al.* Comparative mapping of 9 human chromosome 22q loci in the laboratory mouse. *Hum. Mol. Genet.* **2**, 1245–1252 (1993).
45. Carver, E. A. & Stubbs, L. Zooming in on the human-mouse comparative map: genome conservation re-examined on a high-resolution scale. *Genome Res.* **7**, 1123–1137 (1997).
46. Puech, A. *et al.* Comparative mapping of the human 22q11 chromosomal region and the orthologous region in mice reveals complex changes in gene organization. *Proc. Natl Acad. Sci. USA* **94**, 14608–14613 (1997).
47. Bentley, D. R. Genomic sequence information should be released immediately and freely in the public domain. *Science* **274**, 533–534 (1996).
48. Guyer, M. Statement on the rapid release of genomic DNA sequence. *Genome Res.* **8**, 413 (1998).
49. Dunham, I., Dewar, K., Kim, U.-J. & Ross, M. T. in *Genome Analysis: A Laboratory Manual Series, Volume 3: Cloning Systems* (eds Birren, B. *et al.*) 1–86 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1999).
50. Asakawa, S. *et al.* Human BAC library: construction and rapid screening. *Gene* **191**, 69–79 (1997).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

We thank S. Povey, J. White and H. Wain for the help with gene nomenclature, and M. Adams for making available the sequence trace files of U62317. We thank M. Elharam, H. Jia, L. Lane, R. Morales-Diaz, F. Najjar, P. Pham, R. Rahhal, M. Rao, Y. Tilahun, R. Wayt, H. Wright, E. Nakato, J. L. Schmeits, K. Schooler, J. Wang, M. Asahina, M. Takahashi, H. Harigai, Y. G. Xie, F. Y. Han, S. Swahn, B. Funke, R. K. Pandita, C. Chieffo, D. Michaud and all members of the Sanger Centre past and present for their assistance. This work was supported by grants from the Wellcome Trust, the NIH National Human Genome Research Institute to B.A.R. and to B.S.E., the NSF to B.A.R., the University of Oklahoma, the Fund for the Human Genome Sequencing Project of Japan Science and Technology Corporation, the Fund for the 'Research for the Future' Program from the Japan Society for the Promotion of Science, the UK Medical Research Council, the Medical Research Council of Canada to H.E.M., the Swedish Cancer Foundation, the Swedish Medical Research Council, and a Senior/Established Investigator Award from the Swedish Cancer Foundation to J.P.D.

Correspondence and requests for materials should be addressed to I.D. (e-mail: id1@sanger.ac.uk).

Contributors: I. Dunham¹, A. R. Hunt¹, J. E. Collins¹, R. Bruskewich¹, D. M. Beare¹, M. Clamp¹, L. J. Smink¹, R. Ainscough¹, J. P. Almeida¹, A. Babbage¹, C. Bagguley¹, J. Bailey¹, K. Barlow¹, K. N. Bates¹, O. Beasley¹, C. P. Bird¹, S. Blakey¹, A. M. Bridgeman¹, D. Buck¹, J. Burgess¹, W. D. Burrill¹, J. Burton¹, C. Carder¹, N. P. Carter¹, Y. Chen¹, G. Clark¹, S. M. Clegg¹, V. Copley¹, C. G. Cole¹, R. E. Collier¹, R. E. Connor¹, D. Conroy¹, N. Corby¹, G. J. Coville¹, A. V. Cox¹, J. Davis¹, E. Dawson¹, P. D. Dharmi¹, C. Dockree¹, S. J. Dodsworth¹, R. M. Durbin¹, A. Ellington¹, K. L. Evans¹, J. M. Fey¹, K. Fleming¹, L. French¹, A. A. Garner¹, J. G. R. Gilbert¹, M. E. Goward¹, D. Grafham¹, M. N. Griffiths¹, C. Hall¹, R. Hall¹, G. Hall-Tamlyn¹, R. W. Heathcote¹, S. Ho¹, S. Holmes¹, S. E. Hunt¹, M. C. Jones¹, J. Kershaw¹, A. Kimberley¹, A. King¹, G. K. Laird¹, C. F. Langford¹, M. A. Leversha¹, C. Lloyd¹, I. D. Martyn¹, M. Mashreghi-Mohammadi¹, L. Matthews¹, O. T. McCann¹, J. McClay¹, S. McLaren¹, A. A. McMurray¹, S. A. Milne¹, B. J. Mortimore¹, C. N. Odell¹, R. Pavitt¹, A. V. Pearce¹, D. Pearson¹, B. J. Phillimore¹, S. H. Phillips¹, R. W. Plumb¹, H. Ramsay¹, Y. Ramsey¹, L. Rogers¹, M. T. Ross¹, C. E. Scott¹, H. K. Sehra¹, C. D. Skuce¹, S. Smalley¹, M. L. Smith¹, C. Soderlund¹, L. Spragon¹, C. A. Steward¹, J. E. Sulston¹, R. M. Swann¹, M. Vaudin¹, M. Wall¹, J. M. Wallis¹, M. N. Whiteley¹, D. Willey¹, L. Williams¹, S. Williams¹, H. Williamson¹, T. E. Wilmer¹, L. Wilming¹, C. L. Wright¹, T. Hubbard¹, D. R. Bentley¹, S. Beck¹, J. Rogers¹, N. Shimizu², S. Minoshima², K. Kawasaki², T. Sasaki², S. Asakawa², J. Kudoh², A. Shintani², K. Shibuya², Y. Yoshizaki², N. Aoki², S. Mitsuyama², B. A. Roe², F. Chen², L. Chu², J. Crabtree², S. Deschamps², A. Do², T. Do², A. Dorman², F. Fang², Y. Fu², P. Hu², A. Hua², S. Kenton², H. Lai², H. I. Lao², J. Lewis², S. Lewis², S.-P. Lin², P. Loh², E. Malaj², T. Nguyen², H. Pan², S. Phan², S. Qi², Y. Qian², L. Ray², Q. Ren², S. Shaull², D. Sloan², L. Song², Q. Wang², Y. Wang², Z. Wang², J. White², D. Willingham², H. Wu², Z. Yao², M. Zhan², G. Zhang², S. Chissole², J. Murray², N. Miller², P. Minx², R. Fulton², D. Johnson², G. Bemis², D. Bentley², H. Bradshaw², S. Bourne², M. Cordes², Z. Du², L. Fulton², D. Goela², T. Graves², J. Hawkins², K. Hinds², K. Kemp², P. Latreille², D. Layman², P. Ozersky², T. Rohlfing², P. Scheet², C. Walker², A. Wamsley², P. Wohldmann², K. Pepin², J. Nelson², I. Korf², J. A. Bedell², L. Hillier², E. Mardis², R. Waterston², R. Wilson², B. S. Emanuel², T. Shaikh², H. Kurahashi², S. Saïta², M. L. Budarf², H. E. McDermid², A. Johnson², A. C. C. Wong², B. E. Morrow², L. Edelmann², U. J. Kim², H. Shizuya², M. I. Simon², J. P. Dumanski², M. Peyrard², D. Kedra², E. Seroussi², I. Fransson², I. Tapia², C. E. Bruder², K. P. O'Brien².

Addresses: 1, The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK; 2, Department of Molecular Biology, Keio University School of Medicine, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan; 3, Department of Chemistry and Biochemistry, The University of Oklahoma, 620 Parrington Oval, Room 311, Norman, Oklahoma 73019, USA; 4, Genome Sequencing Centre, Washington University School of Medicine, 4444 Forest Park Blvd, St. Louis, Missouri 63108, USA; 5, Division of Human Genetics and Molecular Biology, The Children's Hospital of Philadelphia and the Department of Pediatrics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104, USA; 6, Department of Biological Sciences, University of Alberta, Edmonton, Alberta, T6G 2E9, Canada; 7, Department of Molecular Genetics, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, New York 10461, USA; 8, California Institute of Technology, Division of Biology, Pasadena, California 91125, USA; 9, Department of Molecular Medicine, Clinical Genetics Unit, Karolinska Hospital, CMM bldg. L8:00, 17176 Stockholm, Sweden

* Present addresses: Division of Virology, Department of Pathology, Tennis Court Road, Cambridge CB2 1QP, UK (A. M. Bridgeman); Lark Technologies, Radwinter Road, Saffron Walden Essex, CB11 3HY, UK (D. Buck, R. E. Connor, S. Ho); Australian Genome Research Facility, Gehrmann Laboratories, University of Queensland, St Lucia QLD 4072, Australia (J. Burgess, J. Davis); Incyte Europe Ltd, 214 Cambridge Science Park, Cambridge CB4 0WA, UK (D. Conroy); Tepel Life Sciences, Innovation Centre, Scotscroft Building, Wilmslow Road, Didsbury, Manchester M20 8RY, UK (S. J. Dodsworth); Department of Brassica & Oilseeds Research, Cambridge Laboratory, John Innes Centre, Norwich, Norfolk, UK (C. Hall); Cancer Genetics Lab, Department of Biochemistry, University of Otago, P.O. Box 56, Dunedin, New Zealand (R. W. Heathcote); Dept. of Zoology, University of Cambridge, Downing Street, CB2 3EJ, UK (M. N. Whiteley); Monsanto, Genomic Team, Mail Zone N3SA, 800 North Lindbergh Boulevard, St. Louis, Missouri 63167, USA (M. Vaudin); Wellcome Trust Centre for Human Genetics, Roosevelt Drive, Oxford OX3 7BN, UK (H. Williamson)



Microtubule dynamics
Temperature sensitivity
Sporulation-induced
Osmo- and cation sensitivity
No phenotype, untested

Multiple phenotypes
Respiration
Haploid invasive mutants
DNA metabolism and protein synthesis
Cell wall biogenesis/maintenance

The temporal requirement for endothelin receptor-B signalling during neural crest development

Myung K. Shin, John M. Levorso, Robert S. Ingram & Shirley M. Tilghman

Howard Hughes Medical Institute and Department of Molecular Biology, Princeton University, Princeton, New Jersey 08544, USA

Endothelin receptor B (EDNRB) is a G-protein-coupled receptor with seven transmembrane domains which is required for the development of melanocytes and enteric neurons. Mice that are homozygous for a null mutation in the *Ednrb* gene are almost completely white and die as juveniles from megacolon. To determine when EDNRB signalling is required during embryogenesis, we have exploited the tetracycline-inducible system to generate strains of mice in which the endogenous *Ednrb* locus is under the control of the tetracycline-dependant transactivators tTA or rtTA. By using this system to express *Ednrb* at different stages of embryogenesis, we have determined that EDNRB is required during a restricted period of neural crest development between embryonic days 10 and 12.5. Moreover, our results imply that EDNRB is required for the migration of both melanoblasts and enteric neuroblasts.

During embryogenesis, neural crest cells arise from the dorsal end of the closing neural tube. These cells are pluripotent, giving rise to a wide range of cell types including neurons and glia of the peripheral nervous system, craniofacial bones and cartilage, smooth muscle cells and melanocytes. Upon their emergence from the neural tube, neural crest cells undergo extensive proliferation as they migrate along distinct pathways to destinations where they differentiate into specific lineages^{1,2}.

Murine coat colour mutants have been useful for identifying genes that are important for the development of neural crest-derived melanocytes^{3,4}. The *Ednrb* gene, encoding a G-protein-coupled receptor with seven transmembrane domains, is deleted in the spontaneous mutant *piebald-lethal* (*Ednrb*^{s-l})⁵. *Ednrb*^{s-l} mutants are almost completely white, except for occasional pigmented areas in the head and rump, and die as juveniles from megacolon, a condition resulting from the inability of neural crest-derived enteric neurons to colonize the distal colon. The ligand for EDNRB is the 21-amino-acid peptide endothelin-3 (EDN3), and mutations in the *Edn3* gene underlie another classic coat colour mutant, *lethal spotting* (*Edn3*^{ls}). These mice closely resemble *Ednrb*^{s-l} mutants, and exhibit severe melanocyte and enteric neuron defects⁶. Mutations in both *EDNRB* and *EDN3* have been described in humans with Hirschsprung's disease and Waardenburg syndrome, who have enteric neuron and pigmentation defects, respectively^{7,8}.

Previous results have indicated that EDNRB signalling may be required at multiple stages of melanocyte development. First, the expression of the melanoblast-specific marker tyrosinase-related protein (TRP-2 or DCT) in *Ednrb*^{s-l} mice is greatly reduced at embryonic day (E) 10.5, implying that EDNRB signalling is required at or before the time of melanoblast proliferation and migration⁹. Second, *in vitro* studies have implicated EDN3 as both a potent mitogen for melanoblast precursor cells, implying an early requirement and a differentiating factor, indicating a possible role for the ligand after birth^{10,11}. Finally, it has been suggested, on the basis of pigmentation differences in *Ednrb*^{s-l} and *Edn3*^{ls} mice, that EDNRB may be required during the intermediate stages of epidermal proliferation at E12.5–13.5 (ref. 12).

For mammalian developmental biologists, the ability to express a gene in a spatially and temporally controlled manner to study its function in the organism has been elusive. In lower eukaryotes, inducible systems such as heat shock and GAL4 are routinely used to achieve conditional expression of genes^{13,14}. Among the many

inducible systems that have been used in mammals¹⁵, the bacterial based tetracycline (tet)-inducible system offers many advantages^{16,17}. These include the low basal expression and high inducibility of the reporter gene, the overall nontoxic effect of tet, and the ability of tet and its derivatives to cross the placenta without affecting the development of embryos^{18,19}.

To date, however, the use of the tet-inducible system in the context of transgenes has met with limited success^{18–22}. The proper spatial expression of the tet-dependent transactivators, tTA and rtTA, has been difficult to accomplish owing to the limited availability of well behaved promoters. In addition, transgene integration is random and, depending upon the site, can result in low or mosaic expression of the activators, as well as leaky expression of the responder genes²³.

To circumvent these problems and to ensure that the *Ednrb* gene is regulated in a spatially appropriate manner, we have integrated the tet-inducible system into the endogenous *Ednrb* locus. We show that this approach leads to tet-induced *Ednrb* expression that can fully rescue the null phenotype and to tet-dependent repression of *Ednrb* expression that recapitulates the null phenotype. By manipulating the expression of the gene during gestation with the tet derivative doxycycline (dox), we have identified a restricted time during which EDNRB signalling is required for the development of both melanoblasts and enteric neurons.

Incorporating the tet system into the *Ednrb* locus

To ensure that the tTA and rtTA proteins were expressed in all *Ednrb*-expressing lineages, we constructed targeting vectors in which the majority of the *Ednrb* first exon and 230 base pairs (bp) of the first intron were replaced with the tTA or rtTA coding region linked to the rabbit β -globin intron/polyadenylation sequences (Fig. 1). The first exon encodes the amino-terminal extracellular domain and the first transmembrane domain of EDNRB; deletion of any one of the membrane-spanning domains of G-protein-coupled seven-transmembrane receptors results in a null allele²⁴. A neomycin-resistance gene flanked by two *loxP* sites was placed downstream of the tTA or rtTA cassettes (Fig. 1a, b). This strategy puts the expression of the transactivator proteins under the control of the endogenous *Ednrb* promoter.

For the responder targeting vector, the region from 550-bp upstream of the putative *Ednrb* promoter to 240-bp downstream of the exon 1/intron 1 junction was replaced with a cassette

containing a tet-responsive promoter that directs the expression of *Ednrb* complementary DNA (*tetO-Ednrb*). This vector contains, in addition to 8.7 kilobases (kb) of *Ednrb* genomic DNA, a hygromycin resistance gene flanked by two *loxP* sites (Fig. 1c). By integrating *tetO-Ednrb* into the *Ednrb* locus, we hoped that the cDNA would be in a chromatin state that would be accessible to the inducer proteins in the appropriate cell types.

The inducer and responder targeting vectors were separately introduced into embryonic stem cells by electroporation, and cell lines that had undergone homologous recombination were identified by Southern analysis with probes that lie outside the targeting vector (Fig. 1a–c; and data not shown). Mice carrying the modified *Ednrb* loci were generated. Initial analysis of the transactivator lines raised the possibility that the presence of the *neo* cassette inhibited expression of *tTA* and *rtTA* from the *Ednrb* promoter. To alleviate this problem, male offspring of the inducer lines were each mated to *Elia-cre* transgenic females, which express *cre* recombinase at high levels in the fertilized egg²⁵, to establish *Ednrb*^{rtTA} and *Ednrb*^{tTA} lines without *neo*.

Ednrb is regulated in utero by dox

To test the efficacy of the tet system, we generated compound heterozygous *Ednrb*^{rtTA}/*Ednrb*^{tetO} offspring, identified with specific

polymerase chain reaction primers designed to distinguish the *Ednrb*, *Ednrb*^{rtTA} and *Ednrb*^{tetO} alleles (data not shown). *rtTA* has been engineered to activate the *tetO* promoter only in the presence of dox (Fig. 1d). As shown in Fig. 1e, two *Ednrb*^{rtTA}/*Ednrb*^{tetO} neonates that developed without dox being given to the mothers were almost completely white, and eventually died of megacolon. Thus the targeted *Ednrb* alleles had disrupted normal *Ednrb* expression, and no leaky expression of *tetO-Ednrb* was evident. *Ednrb* is highly sensitive to gene dosage, and expression of messenger RNA at 12% of wild-type levels leads to pigmentation in the coat²⁶. When dox-containing water was provided to the females for the entire pregnancy, the compound heterozygous offspring exhibited no melanocyte or enteric neuron defects, and no other abnormalities were observed (Fig. 1f).

The complementary *tTA* system, in which dox represses transactivation by *tTA*, was also effective; compound heterozygous *Ednrb*^{tTA}/*Ednrb*^{tetO} mice that developed in the absence of dox were fully pigmented and viable, whereas those raised in the presence of dox displayed the *piebald-lethal* phenotype (Fig. 1d; and data not shown). These results confirmed that dox administered to pregnant females crossed the placenta and could activate or repress the expression of *Ednrb* cDNA from the tet-inducible promoter in the embryos.

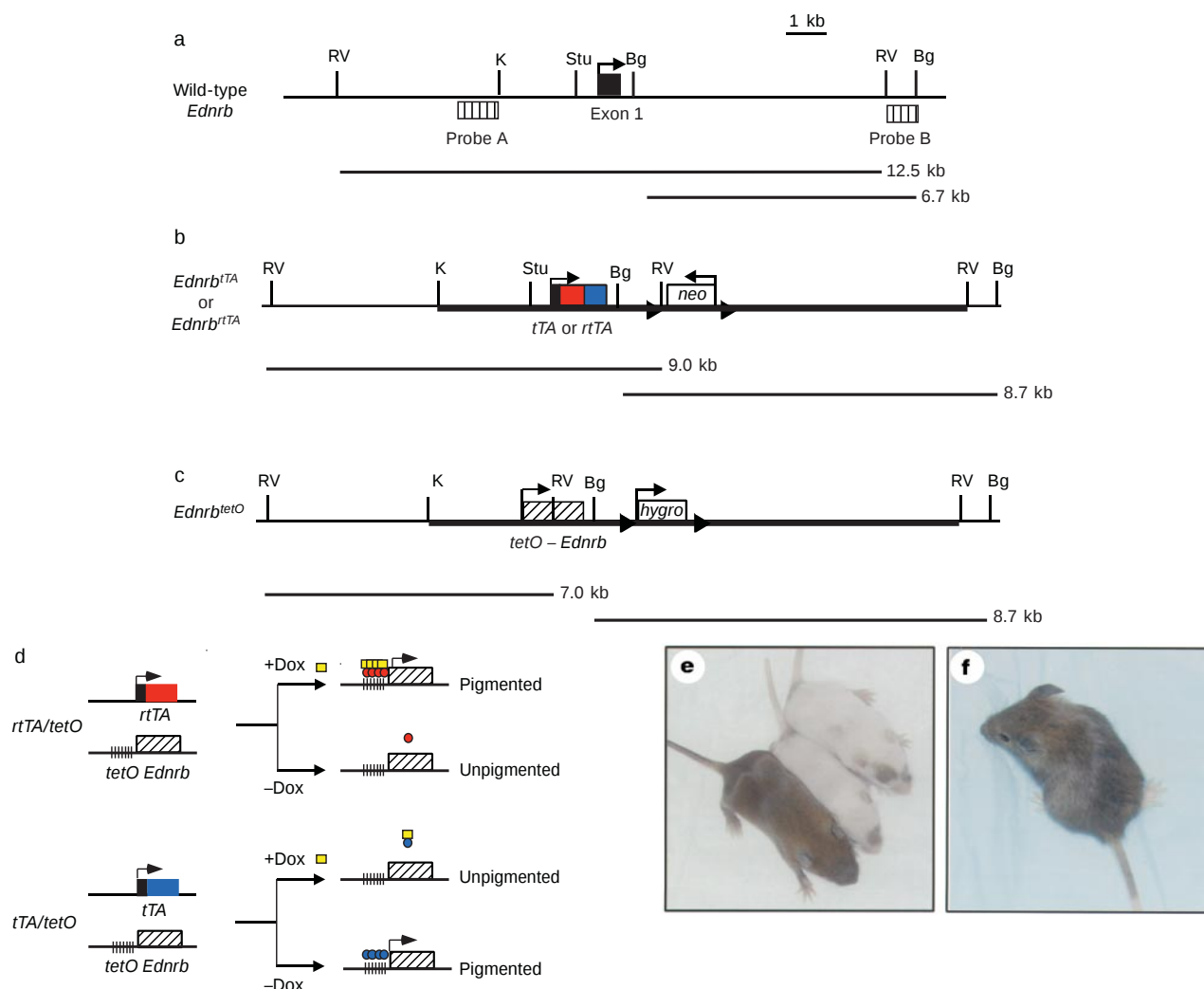


Figure 1 Establishing tet-responsive alleles of *Ednrb*. **a**, The region surrounding the *Ednrb* first exon. **b**, **c**, The targeted *Ednrb* locus containing the transactivators *tTA* or *rtTA* or the tet-responsive *Ednrb* cDNA gene. *LoxP* sites are indicated by filled triangles. The thick lines show the DNA included in the targeting vectors. RV, *EcoRV*; K, *KpnI*; Stu, *StuI*; Bg, *BglII*. **d**, The strategy for regulating *tetO-Ednrb* expression by dox. The *rtTA* and *tTA*

genes and their protein products are indicated in red and blue, respectively. Dox is indicated by a yellow square. **e**, **f**, Unpigmented *tetO-Ednrb*/*rtTA* compound heterozygotes that developed in the absence of dox (**e**) and a pigmented compound heterozygote that developed in the presence of dox (**f**).

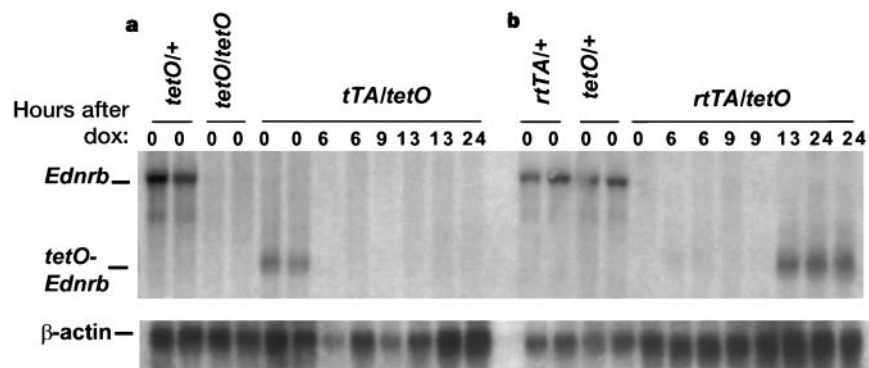


Figure 2 Kinetics of *Ednrb* regulation by dox. **a**, Repression of *tetO-Ednrb* mRNA by dox. *Ednrb*^{tetO}/+ females were mated to *Ednrb*^{tTA}/*Ednrb*^{tetO} males and given dox at E11.5 by oral gavage. Embryos were obtained at the times indicated, and RNA was isolated

and analysed by northern blotting with an *Ednrb* cDNA probe. The blots were stripped and re-probed with a β -actin cDNA. **b**, Induction of *tetO-Ednrb* mRNA by dox. *Ednrb*^{tetO}/+ females were crossed to *Ednrb*^{tTA}/*Ednrb*^{tetO} males, and treated as in **a**.

In vivo kinetics of *Ednrb* activation and repression

To determine the kinetics of *tetO-Ednrb* activation and repression in response to dox *in utero*, mRNA was extracted from individual *Ednrb*^{tTA}/*Ednrb*^{tetO} E11.5 embryos. In the absence of dox, *Ednrb*^{tTA}/*Ednrb*^{tetO} embryos expressed only the 1.8-kb *tetO-Ednrb* mRNA, whereas *Ednrb*^{tetO}/+ littermates expressed only the 4.5-kb endogenous *Ednrb* mRNA, and *Ednrb*^{tetO}/*Ednrb*^{tetO} embryos did not express either *Ednrb* RNA (Fig. 2a). These results confirmed at the molecular level that all targeting events created *Ednrb* null alleles and that no leaky expression occurred from the *tetO* promoter.

Dox was administered to pregnant females by oral gavage on the morning of E11.5, and the animals were maintained on dox water for various times. *Ednrb*^{tTA}/*Ednrb*^{tetO} embryos were identified and examined for *Ednrb* RNA expression. By 6 hours after dox administration, the shortest time examined, the expression of *Ednrb*^{tetO} was undetectable by northern analysis (Fig. 2a).

Ednrb^{tTA}/*Ednrb*^{tetO} embryos did not express any *Ednrb* RNA in the absence of dox. After dox administered by oral gavage, *tetO-Ednrb* mRNA was highly induced within 13 hours (Fig. 2b). The difference in the response kinetics of tTA and rtTA has been described by *in vitro* studies, and implies that the tTA system responds to dox at lower concentrations than the rtTA system²⁷. Despite the kinetic differences, our experiments clearly show that *in utero* activation and repression of *Ednrb* by maternal transmission of dox is rapid and efficient.

The earliest requirement for *Ednrb* function

The latest time when *Ednrb* activation by rtTA can rescue the mutant phenotype was determined by administering the initial dose of dox at different times after fertilization and maintaining the mothers on dox-containing water for the rest of gestation. When dox was administered at or before E9.5, the time when neural crest cells are arising from the neural tube, all *Ednrb*^{tTA}/*Ednrb*^{tetO} offspring were phenotypically wild type (Fig. 3a). As *tetO-Ednrb*

mRNA was induced 9–13 hours after dox treatment (Fig. 2b), we conclude that *Ednrb* is not required before E10.

The compound heterozygotes born to females given dox at E10.5 could be divided into three phenotypic classes: (1) wild type for both melanocytes and enteric ganglia; (2) unpigmented in portions of the head and no megacolon; and (3) unpigmented head and megacolon (Fig. 3b, Table 1). The range of phenotypes could be due to several factors, including the inherent differences in development stages of the embryos within and between litters, differences in genetic background and the nonuniform distribution of dox (and therefore activation of *Ednrb*) within litters. This can be seen in Fig. 3b, where a wild-type *Ednrb*^{tTA}/*Ednrb*^{tetO} pup and a class 3 *Ednrb*^{tTA}/*Ednrb*^{tetO} pup were born in the same litter. These results indicate that the requirement for *Ednrb* signalling begins soon after E10, and that there is a temporal difference in the requirement for *Ednrb* signalling for melanocytes in the head, which are the first to arise, and those in the trunk, which develop later^{12,28–31}.

The litters born to females treated with dox at E11.5 produced two phenotypic groups: (1) amelanocytic, exhibiting the null *piebald-lethal* phenotype; and (2) pigmented in the caudal one- to two-thirds of the body. All progeny died of megacolon within four weeks of birth (Fig. 3c, Table 1). When dox was administered at E12.5 or later, all *Ednrb*^{tTA}/*Ednrb*^{tetO} mice exhibited the *piebald-lethal* phenotype (Fig. 3d). Thus, to rescue both the pigmentation and enteric neuron defects fully, *Ednrb* must be activated between E10 and E11.

The latest requirement for *Ednrb* function

To determine whether *Ednrb* is required later than E10, *Ednrb*^{tetO}/+ females mated to *Ednrb*^{tTA}/*Ednrb*^{tetO} or *Ednrb*^{tTA}/+ males were given dox at different stages during pregnancy to repress *Ednrb* expression. When *tetO-Ednrb* mRNA was repressed by dox administration at E10.5, all *Ednrb*^{tTA}/*Ednrb*^{tetO} compound heterozygotes displayed the full *piebald-lethal* phenotype (Fig. 3e). This result confirmed the

Table 1 Phenotypic analysis of *Ednrb* inducible mutants

Genotype	Duration of <i>Ednrb</i> expression	No. of animals	No. of litters	Pigmentation			Megacolon
				Wild type	Partial pigmentation	Null	
<i>Ednrb</i> ^{tTA} / <i>Ednrb</i> ^{tetO}	E9.5–birth	4	3	4	0	0	0
	E10.5–birth	9	5	3	6 (rostral spotting)	0	4
	E11.5–birth	8	3	0	3 (rostral spotting)	5	8
	E12.5–birth	3	3	0	0	3	3
<i>Ednrb</i> ^{tTA} / <i>Ednrb</i> ^{tetO}	0–E10.5	5	3	0	0	5	5
	0–E11.5	8	6	0	4 (caudal spotting)	4	8
	0–E12.5	3	2	3	0	0	0
	0–E13.5	1	1	1	0	0	0

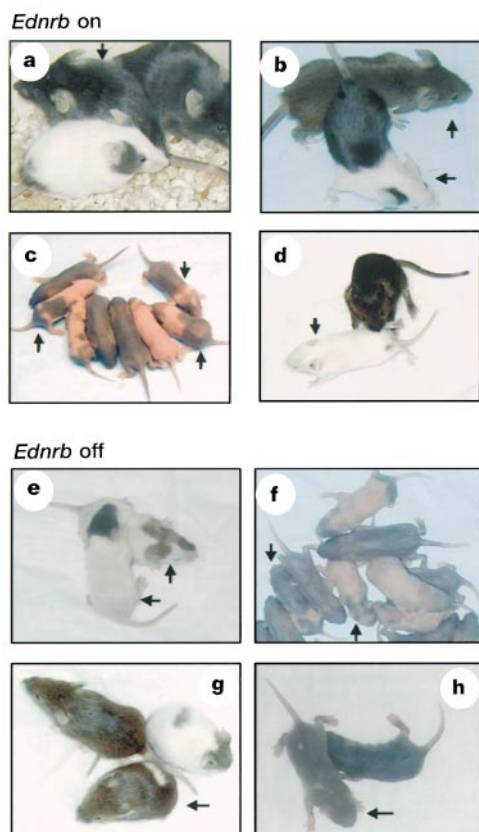


Figure 3 The temporal requirement for *Ednrb* signalling. Pregnant females were given dox at E9.5 (a), E10.5 (b), E11.5 (c) or E12.5 (d), and maintained on dox for the rest of gestation. Arrows identify all *Ednrb*^{tTA}/*Ednrb*^{tetO} heterozygotes. All other mice are *tetO*/*tetO*, *tetO*+/ and *tTA*+/ litter mates. Females were given dox at E10.5 (e), E11.5 (f), E12.5 (g) or E13.5 (h), and maintained on dox for the rest of gestation. Arrows identify all *Ednrb*^{tTA}/*Ednrb*^{tetO} heterozygotes. All other mice were *tetO*/*tetO*, *tetO*+/ and *tTA*+/ litter mates.

conclusion drawn from Fig. 3a–d that *Ednrb* expression is required beyond E10.5. When *Ednrb* expression was inactivated after E11.5, 50% of the compound heterozygotes were *piebald-lethal*. The rest were pigmented in the rostral one- to two-thirds of the body and developed megacolon (Fig. 3f, Table 1). This reciprocal pattern of pigmentation observed in tTA-repressed versus rtTA-activated experiments is consistent with the temporal differences in melanocyte development along the rostral–caudal axis (see Discussion).

The *Ednrb*^{tTA}/*Ednrb*^{tetO} compound heterozygotes obtained when *tetO*–*Ednrb* mRNA was turned off after E12.5 or E13.5 were wild type with the exception of one E12.5 mouse that had an unpigmented area on the lateral part of the torso, as well as a large unpigmented region on the belly (Fig. 3g, h, Table 1). None of these mice developed megacolon (Table 1). As *Ednrb* is completely repressed within 6 hours of administering dox, we conclude that *Ednrb* expression is required before E12.5 for the development of melanoblasts and enteric neuroblasts, but not at later embryonic stages.

Visualization of melanoblasts and enteric neurons

The adult phenotypes of the tet-induced *Ednrb* mice implied a rostral-to-caudal gradient in the rescue of melanoblasts between E10 and E12. To visualize this directly at the time when *Ednrb* is required, we performed *in situ* hybridization on wild-type, mutant and *Ednrb*^{tTA}/*Ednrb*^{tetO} embryos in which dox was administered at E10.5 to activate *Ednrb*. The embryos were hybridized at E11.5 to *Trp2/Dct*, one of the earliest markers of melanoblasts²⁹. As shown

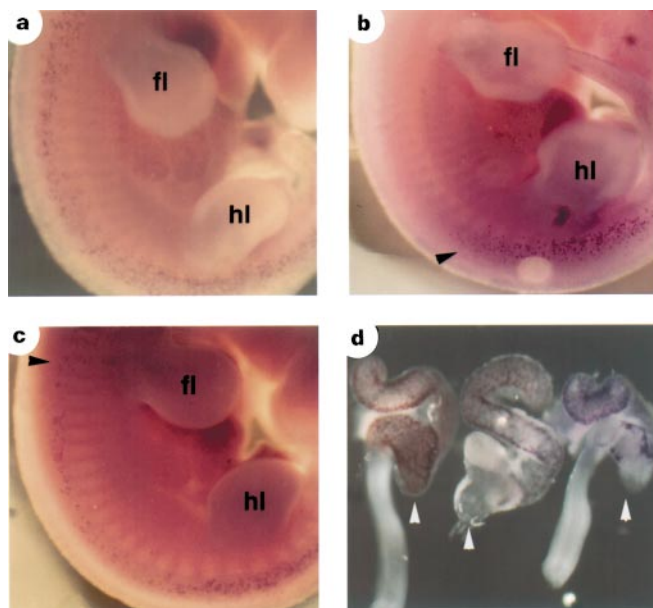


Figure 4 Detection of melanoblasts and enteric neuroblasts by *in situ* hybridization. Pregnant females were given dox at E10.5 and the embryos were removed 24 hours later and hybridized to a digoxigenin-labelled *Trp2/Dct* anti-sense probe. a, Wild-type; b, *Ednrb*^{tetO}/*Ednrb*^{tetO} mutant; c, *Ednrb*^{tTA}/*Ednrb*^{tetO} heterozygote. fl, forelimb bud; hl, hindlimb bud. Arrowheads indicate the most rostral melanoblasts. d, Pregnant females were given dox at E11.5 and the embryos were removed 24 hours later and hybridized to a *c-ret* antisense probe. The guts were dissected from wild-type (left), *Ednrb*^{tTA}/*Ednrb*^{tetO} (middle) and *Ednrb*^{tTA}/*Ednrb*^{tetO} (right) heterozygotes. The arrowheads indicate the position of the junction between the coecum and the colon.

in Fig. 4a, the wild-type embryo exhibited the expected punctate pattern of melanoblasts along the rostral–caudal axis lateral to the neural tube, whereas the mutant displayed staining only in the most caudal region of the embryo. The rescued embryo contained melanoblasts that extended further rostral than those in the mutant, but few melanoblasts were observed anterior to the forelimb buds. This is precisely the pattern one would predict from the adult phenotype (Fig. 3b).

The enteric neuroblasts were also examined using a probe to *c-ret*, an early marker for these cells³². As shown in Fig. 4d, the neuroblasts in the wild-type animals at E12.5 had extended beyond the coecum into the colon, whereas the neuroblasts in the mutant embryo had not proceeded past this junction (data not shown). The *Ednrb*^{tTA}/*Ednrb*^{tetO} embryos that had received dox at E11.5 displayed the mutant phenotype, with no neuroblasts beyond the coecum, very similar to *Ednrb*^{tTA}/*Ednrb*^{tetO} embryos in which the gene had been repressed by dox between E11.5 and E12.5 (Fig. 4d, middle and right). This provides further evidence for a critical window for *Ednrb* signalling between E11.5 and E12.5 for enteric neuroblast development.

Discussion

We have exploited the tet-inducible system to determine when *Ednrb* gene function is required in development. By integrating the complementary tTA and rtTA components into the endogenous *Ednrb* locus, we have achieved robust expression of the *Ednrb* gene from a tet-responsive promoter. Furthermore, our studies demonstrate the utility of the tTA and rtTA transactivators for manipulating the expression of genes with tet *in vivo*.

Neural-crest-derived melanoblasts arise from the dorsal end of the neural tube around E9–E9.5 (Fig. 5). They are presumed to accumulate for 24 hours lateral to the neural tube in a region termed the migrating stage area (MSA). At E10–E10.5, these pre-migratory

melanoblasts begin dispersing dorsolaterally in a temporal rostral–caudal gradient such that by E11 the highest concentrations of migrating melanoblasts are in the cranial and sacral regions. After migrating, melanoblasts enter the epidermis between E12.5 and E13.5, where they proliferate extensively. These cells will complete their differentiation into melanocytes shortly after birth^{12,28–31}.

Our results indicate that *Ednrb* is required between E10 and E12.5, a period corresponding to the developmental stage where melanoblasts are migrating from the MSA (Fig. 5). The fact that the mutant phenotype could be rescued by activating *Ednrb* as late as E10 indicates that *Ednrb* may not be required for the emergence, proliferation or survival of premigratory melanoblasts (Fig. 3a, Table 1). Rather, the restricted temporal requirement for *Ednrb* indicates that its action may be required for the initiation of melanoblast migration and/or for their survival during the developmental transition from premigratory melanoblasts into migratory cells. This conclusion is consistent with the observation that DCT⁺ cells are absent in the trunk of *Ednrb*^{−/−} mice at E10.5 (ref. 9; and Fig. 4b), as experiments have suggested that DCT is detected only in those cells that have left the MSA (in other words, only in migratory melanoblasts)^{31,33}. The full rescue of the mutant phenotype in embryos expressing *Ednrb* from fertilization to E12.5 shows that *Ednrb* is not required for postmigratory epidermal proliferation, survival and differentiation (Fig. 3g, h, Table 1).

A role for *Ednrb* in the initiation of the migratory pathway is supported by the pigmentation patterns we observed in mice where *Ednrb* was activated or repressed between E10.5 and E11.5. The rescue of caudal but not rostral cells in some *Ednrb*^{rtTA}/*Ednrb*^{tetO} compound heterozygotes that developed until E10.5 or E11.5 in the absence of dox indicates that caudal melanoblasts require *Ednrb* function later than do the rostral cells. This result is consistent with the observation that caudal cells emerge later than rostral cells (Figs 3b, c and 4c). The absence of the most rostral pigment cells could be

due to the loss of melanoblasts in the MSA, which would explain why later *Ednrb* expression fails to rescue the cells, and/or the inability of the MSA melanoblasts to enter the migratory pathway at later stages, possibly because the environment is no longer permissive. In the reciprocal experiment, where *Ednrb* is turned off at E11.5, rostral cells were rescued more efficiently than caudal cells (Fig. 5b, Table 1). This pattern is consistent with rostral melanoblasts having initiated their migration before the cessation of *Ednrb* expression, whereas caudal cells were still in the MSA. Rostral melanoblasts in the epidermis have been observed as early as E11 (ref. 34).

Finally, these studies show that *Ednrb* is also required between E10.5 and E12.5 for the development of enteric neuroblasts (Table 1). This conclusion is supported by observations using a *lacZ*-tagged *Ednrb* null allele which showed that the progress of migrating enteric neuroblasts is delayed by E11 in mutant embryos (M.K.S., unpublished data). Likewise, the migration of the enteric neurons is displayed in *Ednrb*-null *spotting lethal* rats³⁵. However, our observations differ from previous studies that concluded that differences between wild-type and mutant mice are not observed until E12 (ref. 36). The discrepancy may be due to differences in the genetic backgrounds or to the fact that the earlier study used transgenic mice marked with a *dopamine-β-hydroxylase lacZ* gene to identify enteric neuroblasts and this marker may not be expressed in all *Ednrb*-expressing cells. On the other hand, our results are consistent with a model^{37,38} proposing that EDN3 may be required to maintain a sufficient pool of enteric neuronal precursors in a migratory undifferentiated state. According to this model, in the absence of *Ednrb* signalling migrating enteric neuroblasts differentiate prematurely and become postmitotic, thereby failing to colonize the lower bowel. The finding that *Edn3* is expressed at the highest levels in the coecum is consistent with this model³⁹. Thus it could be argued that our study has uncovered a unifying logic for the function of *Ednrb* in regulating migration of both melanoblasts and enteric neuroblasts.

Despite our successful use of the tet-inducible system, we note several limitations. First, the level of *tetO-Ednrb* mRNA in the absence of dox in *tTA* animals or in the presence of dox in *rtTA* mice, although sufficient to rescue the mutant phenotype, is lower than that of endogenous *Ednrb* mRNA in heterozygotes (~60%; Fig. 2). By improving splicing and codon usage, the production of *tTA* and *rtTA* have been greatly improved in mammalian cells²³. Second, although the *tetO* promoter was repressed efficiently in *rtTA* mice in the absence of dox, or in *tTA* mice in the presence of the drug (Fig. 2), this was not the case when we deleted the *Pgk1-hygro* cassette from the responder line. Instead, we obtained partially pigmented *Ednrb*^{rtTA}/*Ednrb*^{tetO} mice that resembled the hypomorphic *Ednrb*^s allele³ in the absence of dox (data not shown). Thus, it appears that the *Pgk-1* promoter fortuitously inhibited the basal *tetO* promoter.

Despite these limitations, the integration of the responder and activators into the endogenous locus by homologous recombination alleviated many problems associated with previous studies using transgenic mice. The ability of dox to cross the placenta and rapidly regulate gene expression *in utero*, coupled with the considerable flexibility that the tet system affords to investigators, makes this system very useful for studying mammalian development. The *tTA* and *rtTA* mice that we generated will be generally useful for studying other genes implicated in melanocyte and enteric neuron development, as well as genes that function in other lineages where *Ednrb* is expressed. □

Methods

Targeting vectors and embryonic stem cells

Ednrb genomic DNA was isolated from a 129Sv/J λFIXII library (Stratagene). For the transactivator constructs, the forward primer 5'-CTGCAGCACCATTGTCTAGAT TAGA-TAAAGT-3' and the reverse primer 5'-GGTACCGCCACCTTGTTCATGG CAGC-3' were used to PCR amplify the *tTA* and *rtTA* cassettes from plasmids pUHG15-1 (ref. 18)

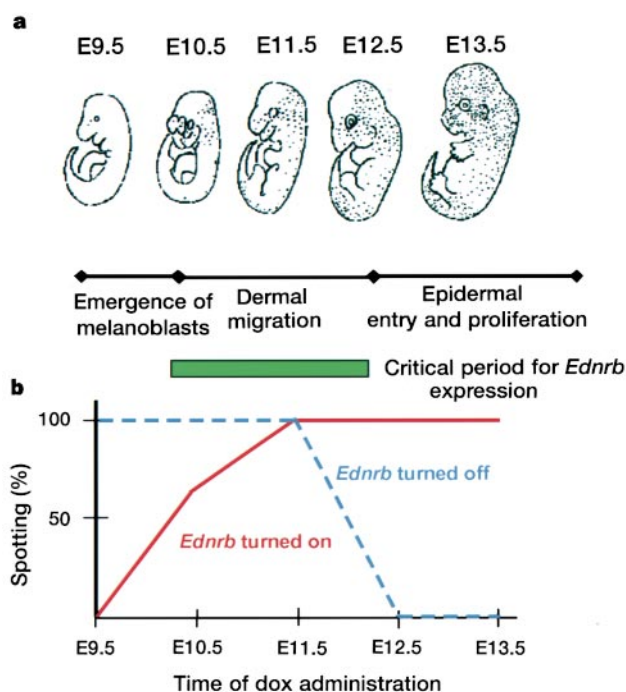


Figure 5 *Ednrb* signalling is required during melanoblast migration. **a**, The appearance of c-kit⁺ and TRP2⁺ melanoblasts in mouse embryos at different stages of embryogenesis is adapted from Yoshida *et al.*¹². **b**, The degree of white spotting in *Ednrb*^{rtTA}/*Ednrb*^{tetO} heterozygotes (blue) and *Ednrb*^{rtTA}/*Ednrb*^{tetO} heterozygotes (red) treated with dox for the times indicated. The critical period when *Ednrb* signalling is required is indicated by the green box.

and pUHG17-1 (ref. 17), respectively. The targeting vector contained PGK-neo^R flanked by two *loxP* sites, a 3.3-kb genomic fragment that contains 200 bp of the untranslated region but not the *Ednrb* ATG initiation codon and a 6-kb *Bg*II-*Eco*RV 3' flanking fragment. These constructs were transfected into E14.1 embryonic stem (ES) cells⁴⁰, and positive clones were selected with 200 µg ml⁻¹ active G418 as described⁴¹. For the responder construct, a 1.4-kb *Ednrb* cDNA containing the entire coding region of the protein²⁶ was subcloned downstream of the tetO promoter of pUHG10-3 (a gift from L. Hennighausen). The targeting vector contained PGK-hygro^R flanked by *loxP* sites, a 2.7-kb *Kpn*I-*Stu*I 5' genomic fragment and a 6-kb *Bg*II-*Eco*RV 3' genomic fragment. The targeting vector was transfected into C17 ES cells⁴², and positive clones were selected with 150 µg ml⁻¹ active hygromycin as described⁴³. ES cells were injected into C57BL/6J blastocysts and germline mice were generated as described⁴³.

Genotyping *Ednrb* alleles

The three targeted *Ednrb* alleles were distinguished by PCR of tail or yolk sac DNA using the following primers: *Ednrb*5', 5'-CCAGACTGAAAACAGCAGAGCGGC-3'; *Ednrb* 2, 5'-GGTCTCCAGAGCCAGACTGGCGATC-3'; CMV407, 5'-GGCGTGTACGGTGG-GAGG-3'; tTA416, 5'-GCATTAGAGCTGCTTAATGAGG-3'; tTA415, 5'-TCTTGATCTTCCAATACGCAACC-3'. The *Ednrb*5' and *Ednrb*2 primers produce a 495-bp wild-type *Ednrb* exon1 fragment. The CMV407 (ref. 18) and *Ednrb*2 primers identify a 610-bp *tetO*-*Ednrb* fragment. Primers tTA416 and tTA415 amplify 410-bp tTA or rtTA fragments. These alleles were further distinguished by *Nru*I and *Hind*III restriction fragment length polymorphisms within the PCR products. The PCR reactions were performed under the following conditions: 94 °C for 30 s, 57 °C for 45 s, 72 °C for 60 s for 35 cycles.

The PGK-neo cassette was deleted from *Ednrb*^{tTA}/+ and *Ednrb*^{rtTA}/+ mice by mating heterozygous males from each line to E1a-cre females as described²⁵.

Doxycycline administration

For timed pregnancies, the morning in which the vaginal plug was observed was E0.5. An initial dose of 200 µl of 1 mg ml⁻¹ dox (Sigma) was administered to pregnant females by oral gavage between 9 and 10 AM. Subsequently, the females were maintained on 1 mg ml⁻¹ dox water for the duration of the pregnancies. The megacolon phenotype was determined by observing a gross distension of the colon post-mortem in affected animals as described⁵.

RNA analysis

Total RNA was isolated from E11-E12 embryos by Trizol extraction (Gibco-BRL). PolyA+ RNA was isolated by the Micro FastTrack Kit (Invitrogen) and analysed by agarose gel electrophoresis and northern blotting²⁶.

In situ hybridization

Whole-mount *in situ* hybridization was performed on embryos as described⁴⁴. To detect melanoblasts, a *Trp2/Dct*-specific riboprobe²⁹ was hybridized to *Ednrb*^{tTA}/*Ednrb*^{tetO} E11.5 embryos that had been given dox at E10.5. To detect enteric neuroblasts, a *c-RET*-specific riboprobe³² was hybridized to *Ednrb*^{tTA}/*Ednrb*^{tetO} and *Ednrb*^{tTA}/*Ednrb*^{tetO} E12.5 embryos that had been given dox at E11.5.

Received 15 June; accepted 12 October 1999.

1. Le Douarin, N. M. *The Neural Crest* (Cambridge Univ. Press, Cambridge, 1982).
2. Anderson, D. J. *et al.* Cell lineage determination and the control of neuronal identity in the neural crest. *Cold Spring Harb. Symp. Quant. Biol.* **62**, 493-504 (1997).
3. Silvers, W. K. *The Coat Colors of Mice: A Model for Mammalian Gene Action and Interaction* (Springer, New York, 1979).
4. Barsh, G. S. The genetics of pigmentation: from fancy genes to complex traits. *Trends Genet.* **12**, 299-305 (1996).
5. Hosoda, K. *et al.* Targeted and natural (piebald-lethal) mutations of endothelin-B receptor gene produce megacolon associated with spotted coat color in mice. *Cell* **79**, 1267-1276 (1994).
6. Baynash, A. G. *et al.* Interaction of endothelin-3 with endothelin-B receptor is essential for development of epidermal melanocytes and enteric neurons. *Cell* **79**, 1277-1285 (1994).
7. Bedell, M. A., Largaespada, D. A., Jenkins, N. A. & Copeland, N. G. Mouse models of human disease. Part II: recent progress and future directions. *Genes Dev.* **11**, 11-43 (1997).
8. Jackson, I. J. Homologous pigmentation mutations in human, mouse and other model organisms. *Hum. Mol. Genet.* **6**, 1613-1624 (1997).
9. Pavan, W. J. & Tilghman, S. M. Piebald lethal (sl) acts early to disrupt the development of neural crest-derived melanocytes. *Proc. Natl Acad. Sci. USA* **91**, 7159-7163 (1994).
10. Lahav, R., Ziller, C., Dupin, E. & Le Douarin, N. M. Endothelin 3 promotes neural crest cell proliferation and mediates a vast increase in melanocyte number in culture. *Proc. Natl Acad. Sci. USA* **93**, 3892-3892 (1996).
11. Reid, K. *et al.* Multiple roles for endothelin in melanocyte development: regulation of progenitor number and stimulation of differentiation. *Development* **122**, 3911-3919 (1996).
12. Yoshida, H., Kunisada, T., Kusakabe, M., Nishikawa, S. & Nishikawa, S. I. Distinct stages of melanocyte differentiation revealed by analysis of nonuniform pigmentation patterns. *Development* **122**, 1207-1214 (1996).
13. Schneuwly, S., Klemenz, R. & Gehring, W. J. Redesigning the body plan of *Drosophila* by ectopic expression of the homeotic gene Antennapedia. *Nature* **325**, 816-818 (1987).

14. Fischer, J. A., Giniger, E., Maniatis, T. & Ptashne, M. GAL4 activates transcription in *Drosophila*. *Nature* **332**, 853-856 (1988).
15. Gingrich, J. R. & Roder, J. Inducible gene expression in the nervous system of transgenic mice. *Annu. Rev. Neurosci.* **21**, 377-405 (1998).
16. Gossen, M. & Bujard, H. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc. Natl Acad. Sci. USA* **89**, 5547-5551 (1992).
17. Gossen, M. *et al.* Transcriptional activation by tetracyclines in mammalian cells. *Science* **268**, 1766-1769 (1995).
18. Furth, P. A. *et al.* Temporal control of gene expression in transgenic mice by a tetracycline-responsive promoter. *Proc. Natl Acad. Sci. USA* **91**, 9302-9306 (1994).
19. Kistner, A. *et al.* Doxycycline-mediated quantitative and tissue-specific control of gene expression in transgenic mice. *Proc. Natl Acad. Sci. USA* **93**, 10933-10938 (1996).
20. Mayford, M. *et al.* Control of memory formation through regulated expression of a CaMKII transgene. *Science* **274**, 1678-1683 (1996).
21. Ewald, D. *et al.* Time-sensitive reversal of hyperplasia in transgenic mice expressing SV40 T antigen. *Science* **273**, 1384-1386 (1996).
22. Mansuy, I. M. *et al.* Inducible and reversible gene expression with the rtTA system for the study of memory. *Neuron* **21**, 257-265 (1998).
23. Rossant, J. & McMahon, A. "Cre"-ating mouse mutants—a meeting review on conditional mouse genetics. *Genes Dev.* **13**, 142-145 (1999).
24. Dixon, R. A. *et al.* Structural features required for ligand binding to the beta-adrenergic receptor. *EMBO J.* **6**, 3269-3275 (1987).
25. Lakso, M. *et al.* Efficient *in vivo* manipulation of mouse genomic sequences at the zygote stage. *Proc. Natl Acad. Sci. USA* **93**, 5860-5865 (1996).
26. Shin, M. K., Russell, L. B. & Tilghman, S. M. Molecular characterization of four induced alleles at the *Ednrb* locus. *Proc. Natl Acad. Sci. USA* **94**, 13105-13110 (1997).
27. Baron, U. *et al.* Generation of conditional mutants in higher eukaryotes by switching between the expression of two genes. *Proc. Natl Acad. Sci. USA* **96**, 1013-1018 (1999).
28. Serbedzija, G. N., Fraser, S. E. & Bronner-Fraser, M. Pathways of trunk neural crest cell migration in the mouse embryo as revealed by vital dye labelling. *Development* **108**, 605-612 (1990).
29. Steel, K. P., Davidson, D. R. & Jackson, I. J. TRP-2/DT, a new early melanoblast marker, shows that steel growth factor (c-kit ligand) is a survival factor. *Development* **115**, 1111-1119 (1992).
30. Wehrle-Haller, B. & Weston, J. A. Soluble and cell-bound forms of steel factor activity play distinct roles in melanocyte precursor dispersal and survival on the lateral neural crest migration pathway. *Development* **121**, 731-742 (1995).
31. Nakayama, A. *et al.* Mutations in microphthalmia, the mouse homolog of the human deafness gene MITF, affect neuroepithelial and neural crest-derived melanocytes differently. *Mech. Dev.* **70**, 155-166 (1998).
32. Pachnis, V., Mankoo, B. & Costantini, F. Expression of the c-ret proto-oncogene during mouse embryogenesis. *Development* **119**, 1005-1017 (1993).
33. Mackenzie, M. A., Jordan, S. A., Budd, P. S. & Jackson, I. J. Activation of the receptor tyrosine kinase Kit is required for the proliferation of melanoblasts in the mouse embryo. *Dev. Biol.* **192**, 99-107 (1997).
34. Mayer, T. C. The migratory pathway of neural crest cells into the skin of mouse embryos. *Dev. Biol.* **34**, 39-46 (1973).
35. Newgreen, D. F. & Hartley, L. Extracellular matrix and adhesive molecules in the early development of the gut and its innervation in normal and spotting lethal rat embryos. *Acta Anat. (Basel)* **154**, 243-260 (1995).
36. Kapur, R. P., Sweetser, D. A., Doggett, B., Siebert, J. R. & Palmiter, R. D. Intercellular signals downstream of endothelin receptor-B mediate colonization of the large intestine by enteric neuroblasts. *Development* **121**, 3787-3795 (1995).
37. Hearn, C. J., Murphy, M. & Newgreen, D. GDNF and ET-3 differentially modulate the numbers of avian enteric neural crest cells and enteric neurons *in vitro*. *Dev. Biol.* **197**, 93-105 (1998).
38. Wu, J. J., Chen, J. X., Rothman, T. P. & Gershon, M. D. Inhibition of *in vitro* enteric neuronal development by endothelin-3: mediation by endothelin B receptors. *Development* **126**, 1161-1173 (1999).
39. Leibel, M. A. *et al.* Expression of endothelin 3 by mesenchymal cells of embryonic mouse caecum. *Gut* **44**, 246-252 (1999).
40. Kuhn, R., Rajewsky, K. & Muller, W. Generation and analysis of interleukin-4 deficient mice. *Science* **254**, 707-710 (1991).
41. Jones, B. K., Levorse, J. M. & Tilghman, S. M. Igf2 imprinting does not require its own DNA methylation or H19 RNA. *Genes Dev.* **12**, 2200-2207 (1998).
42. Swiatek, P. J. & Gridley, T. Perinatal lethality and defects in hindbrain development in mice homozygous for a targeted mutation of the zinc finger gene *Krox20*. *Genes Dev.* **7**, 2071-2084 (1993).
43. Webber, A. L., Ingram, R. S., Levorse, J. M. & Tilghman, S. M. Location of enhancers is essential for the imprinting of H19 and Igf2 genes. *Nature* **391**, 711-715 (1998).
44. Wilkinson, D. G. & Nieto, M. A. Detection of messenger RNA by *in situ* hybridization to tissue sections and whole mounts. *Methods Enzymol.* **225**, 361-373 (1993).

Acknowledgements

We thank L. Hennighausen and H. Bujard for providing constructs; H. Westphal for providing E1a-cre transgenic mice; the members of our laboratory for helpful discussions; and M. Cleary for critical comments on the manuscript. M.K.S. was a Jane Coffin Childs Postdoctoral Fellow and S.M.T. is an Investigator of the Howard Hughes Medical Institute. R.S.I. isolated and characterized the *Ednrb* genomic clones and J.M.L. injected the targeted ES cells into blastocysts. M.K.S. and S.M.T. were responsible for the design of the experiment and the writing of the manuscript, and M.K.S. for the generation of the targeting vectors, targeting of ES cells and analysing the mice.

Correspondence and requests for materials should be addressed to S.M.T.

Discovery of a massive equatorial torus in the η Carinae stellar system

P. W. Morris[†], L. B. F. M. Waters[‡], M. J. Barlow[§], T. Lim^{||}, A. de Koter^{*}, R. H. M. Voors[¶], P. Cox[#], Th. de Graauw[☆], Th. Henning^{**}, S. Hony^{††}, H. J. G. L. M. Lamers[†], H. Mutschke^{**} & N. R. Trams^{††}

^{*} Astronomical Institute "Anton Pannekoek", University of Amsterdam, Kruislaan 403, NL-1098 SJ Amsterdam, The Netherlands

[†] SRON Laboratory for Space Research, Sorbonnelaan 2, NL-3584 CA Utrecht, The Netherlands

[‡] Instituut voor Sterrenkunde, Katholieke Universiteit Leuven, Celestijnenlaan 200B, B-3001 Heverlee, Belgium

[§] University College London, Gower Street, London WC1E 6BT, UK

^{||} Rutherford Appleton Laboratories, Chilton, Didcot OX11 0QX, UK

[¶] Astronomical Institute, Utrecht University, Princetonplein 5, NL-3508 TA Utrecht, The Netherlands

[#] Institut d'Astrophysique Spatiale, Bâtiment 121, Université de Paris XI, F-91405 Orsay, France

[☆] SRON Laboratory for Space Research Groningen, PO Box 800, NL-9700 AV Groningen, The Netherlands

^{**} Astrophysical Institute and University Observatory, Friedrich Schiller University, Schillergäßchen 3, D-07745 Jena, Germany

^{††} INTEGRAL Science Operations Centre, Astrophysics Division of ESA, ESTEC, PO Box 299, NL-2200 AG Noordwijk, The Netherlands

The enigmatic object η Carinae is believed to represent an important, but short-lived, unstable phase in the life of the most massive stars, occurring shortly before they explode as supernovae or collapse directly to black holes. The putative binary^{1,2} system believed to constitute η Carinae survived an outburst in the previous century that lasted 20 years; and which created a nebula with pronounced bipolar lobes that together contain about 2.5 solar masses of material. The nebula also exhibits an equatorial 'waist' containing about 0.5 solar masses³. The physical mechanisms responsible for the outburst and the bipolar geometry are not understood. Here we report infrared observations (spectroscopy and imaging) that reveal the presence of about 15 solar masses of material, located in an equatorial torus. The massive torus may have been created through highly non-conservative mass transfer, which removed the entire envelope of one of the stars, leaving an unstable core that erupted in the nineteenth century. The collision of the erupted material with the pre-existing torus provides a natural explanation for the bipolar shape of the nebula.

The stellar system η Car may consist of two very massive (50–80 $M_{\odot}$, where $M_{\odot}$ is the solar mass) stars in a wide, 5.5-year orbit^{1,2}. The binary is surrounded by a highly bipolar, dusty nebula containing 2–3 $M_{\odot}$ of material^{3–5}. This bipolar nebula was ejected by one of the stars in the system over a 20-year period starting in 1843. A second major eruption occurred around 1890. No other stellar objects (apart from supernovae) are known to have shown such extreme mass loss. The origin of the two outbursts and the highly bipolar shape of the present-day nebula are not known. Models for the present-day shape of the nebula involve the symmetric ejection of 2–3 $M_{\odot}$ of material into a massive pre-existing equatorial disk or torus^{6,7}. This massive pre-existing disk should have contained at least 2 $M_{\odot}$ of material in order for it to have survived wind ablation⁷. Theoretically, such a massive disk could have been created by a rapidly rotating massive star⁸; but although many studies have speculated on the presence of a massive equatorial torus^{9–12}, observations to date have revealed evidence for only about 0.5 $M_{\odot}$ of material in the equatorial waist^{13–15}. Others¹⁶ have speculated on the creation of a temporary massive torus following the 1843 outburst, with the aim of explaining the 1890 event. None of these scenarios address the origin of the first giant outburst.

Studies of the radial velocities of the ejecta show the presence of slowly moving material in the equatorial regions of the nebula, ejected 1,000–2,000 years (kinematical age) before the 1843 event¹³. Condensations detected with imaging at 12.5 μm have been proposed to be loop-like structures¹⁷, while other workers^{9–12} had already speculated that material from the giant eruption swept into a pre-existing dense toroidal ring of 1–2 $M_{\odot}$ in total mass; this supposed ring was confined to the inner arcsecond and was probably destroyed by the eruption. Here we report new infrared images and spectra of the nebula surrounding η Car, revealing the mass distribution of dust in the equatorial regions. We find evidence for a massive torus which may have deflected an essentially spherical flow of material ejected from one of the stars during the 1843 outburst, forming the bipolar nebula.

An infrared spectrum of η Car (Fig. 1) was taken by the Infrared Space Observatory (ISO)¹⁸, and can be well fitted by a model with three principal thermal components: a hot power-law component (see below) between 2 and 8 μm , and two modified black-body components with temperatures, T_{dust} of 190 ± 10 and 110 ± 10 K, respectively. Previous studies report^{14,15} black-body components of 400 K and 200–250 K. Our observations show that the 400 K component is better represented by a power-law $F_{\nu} \propto \nu^{-3.0}$ (where F_{ν} is the flux density in $\text{ergs cm}^{-2} \text{s}^{-1} \text{Hz}^{-1}$ at frequency ν), consisting of thermal emission at temperatures between 1,500 and 400 K. The 200–250 K component is similar to the warmer (190 K) black-body component in the ISO spectrum. The cooler (110 K) component evident in the ISO spectrum, which contains by far the most of the circumstellar mass, has not been previously identified.

The temperature of the cool component is constrained by the

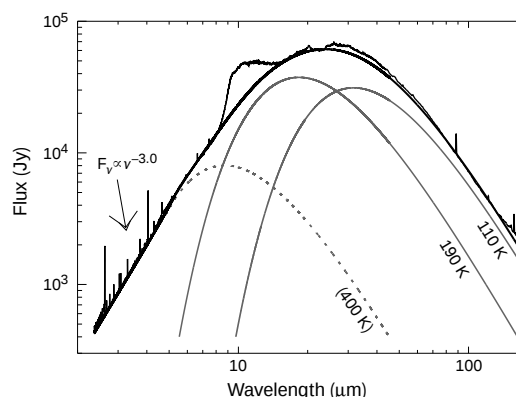


Figure 1 Infrared spectrum of η Car. η Car was observed with the short wavelength spectrometer (SWS)²¹ and the long wavelength spectrometer (LWS)²² on board the Infrared Space Observatory (ISO)¹⁸. Full grating scans were obtained with both instruments in January 1996. A detailed description of the spectral and photometric calibrations and data reduction will be given elsewhere. The 35–175 μm fluxes are significantly higher than found in analyses of photometry obtained in the 1970s with the Kuiper Airborne Observatory (KAO) over 5 passbands between 35 and 175 μm (ref. 23). We have resolved these differences as a matter of calibration of the older photometry, based on observations of Uranus as the photometric standard. With new predictions of the brightness temperatures (provided by E. Lellouch) based on SWS and LWS observations of Uranus, we find the KAO photometry of η Car to agree reasonably well with the ISO data. Integrating under the ISO fluxes, we obtain an infrared luminosity $L_{\text{IR}} = 10^{6.8} L_{\odot}$ for a distance of 2.3 kpc, which agrees well with previous estimates¹⁴. (Here $L_{\odot}$ is the Sun's luminosity.) The total luminosity is $L_{\star} \approx 10^{6.7} L_{\odot}$, assuming that 20% of the ultraviolet and optical radiation is unscattered. The contribution of continuum radiation from the stellar system to the observed fluxes is negligible. Temperatures of the 190 and 110 K thermal components are determined from fitting modified Planck functions to the ISO spectrum, and have uncertainties of ± 10 K. The modified black bodies have the form $\nu^{\beta} B_{\nu}(T_{\text{dust}})$, where β is determined by grain emissivity properties. The 40–200 μm portion of the ISO spectra yields a best-fit value of $\beta = 1.22$.

peak intensity in the ISO spectrum near $25\ \mu\text{m}$, and also by the knowledge of the $\sim 200\ \text{K}$ component associated with the bipolar lobes that contain about $2\ M_{\odot}$ of material^{14,15}. More specifically, we find that the highest-temperature black body that does not exceed the observed fluxes at $25\ \mu\text{m}$ has $T_{\text{dust}} = 130\ \text{K}$. Then the most acceptable fit to the overall distribution requires that the warm component has T_{dust} of $250\ \text{K}$, but is substantially reduced in its contribution to the total flux. Together, these would lead to a negligible amount of material in the lobes, inconsistent with results from imaging over several infrared passbands out to $12.8\ \mu\text{m}$ (ref. 15).

The total dust mass within each component can be estimated from the relation $M_{\text{dust}} = F_{\nu} d^2 / B_{\nu}(T_{\text{dust}}) \kappa_{\nu}$, where $d = 2.3\ \text{kpc}$ is the distance to $\eta\ \text{Car}$ (ref. 3), B_{ν} is the Planck function evaluated at temperature T_{dust} , and κ_{ν} describes the dust grain absorption and scattering efficiencies weighted by grain size and density. Assuming a silicate dust composition, we find dust masses ($M_{\odot}$) of 3.0×10^{-4} , 0.020 ± 0.004 and 0.15 ± 0.05 for the hot, $190 \pm 10\ \text{K}$ and $110 \pm 10\ \text{K}$ dust, respectively. The use in the model of large ($1\text{--}10\ \mu\text{m}$) silicate grains, which are efficient scatterers of radiation from the central source, provides a lower estimate to the dust masses. The power-law component and lobe mass estimates are consistent with previous studies^{14,15}. The dust mass of the cool component dominates, increasing the total dust mass in the equatorial region compared to previous studies by a factor of about 20–40.

Adopting a canonical gas-to-dust mass ratio of 100, as in other

mid-infrared studies of $\eta\ \text{Car}$, implies a total mass of $15\ M_{\odot}$. However, this adopted ratio is uncertain by a factor of two (with the tendency towards the lower value in the Galactic Centre and higher values in the Magellanic clouds), so that the overall uncertainty on the total mass of $15\ M_{\odot}$ for the cool component is estimated to be a factor 2.5. This does not take into account the assumptions about grain emissivity, which were made to give a lowest limit to the dust mass. Also, the inclusion of additional components at temperatures between 110 and $190\ \text{K}$ would have the effect of decreasing the dust mass in the 110 and $190\ \text{K}$ components. The warmer-temperature components would be much more strongly diluted, due to the increasing competition from thermal peaks towards shorter wavelengths. But if we constrain the dust mass in the lobes to be at least $0.01\ M_{\odot}$ (to be marginally consistent with results from $2\text{--}13\ \mu\text{m}$ imaging¹⁵), then the contribution from the intermediate temperatures to the total flux (or mass) will have very little effect on the mass estimate for the cool component.

Ground-based infrared images can be used to locate the three spectral components that are present in the ISO spectra. Images taken at $4.8\ \mu\text{m}$ show that the flux is dominated by an unresolved central component coincident with the position of the binary system¹⁵. We conclude that the power-law component originates from this compact, unresolved ($<1\ \text{arcsec}$) central region. A power-law energy distribution is characteristic of optically thick dust, with a power-law radial temperature and density distribution. Near $8\ \mu\text{m}$ is where the material becomes optically thin. The optically thick material cannot be spherically distributed around the central object, as optical light scattered off dust grains in the lobes of the nebula show that this material has a relatively unobscured view of the central source. This strongly suggests that the matter is distributed in a disk geometry, approximately perpendicular to the lobes. The power-law emitting region is spatially unresolved in our imaging.

Comparison of the nebular brightness at 4 and $12\ \mu\text{m}$ shows that the $190\ \text{K}$ component is located in the polar lobes¹⁵. We have obtained very high quality 10-- and $17\text{--}\mu\text{m}$ ground-based images to find the dust distribution of the coolest material in the nebula. The $17\text{--}\mu\text{m}$ image with $10\text{--}\mu\text{m}$ contours overlaid (Fig. 2) shows that the central unresolved component is very faint at $17\ \mu\text{m}$, consistent with its high temperature. The brightness distribution at $17\ \mu\text{m}$ is dominated by two bright spots in the equatorial 'waist', separated by $\sim 4\ \text{arcsec}$. These spots (in particular the south-west region of the waist) are less prominent at $10\ \mu\text{m}$, indicating that they contain cool dust. They were not noticed in a previously published $17\text{--}\mu\text{m}$ image as a probable result of the quoted difficult observing conditions¹⁷. The central regions in our 10-- and $17\text{--}\mu\text{m}$ images are consistent with emission from a limb-brightened torus of cold dust, which we identify with the $110\ \text{K}$ component. A massive torus provides the only geometry that can shield the grains from the stellar radiation field, resulting in an equilibrium dust temperature of $110\ \text{K}$ so close to the star.

The discovery of an equatorial torus with a mass which is unprecedented for evolved massive stars allows us to re-evaluate the history of the $\eta\ \text{Car}$ system, including the formation of the bipolar nebula and the possible origin of the nineteenth-century outburst itself. Hydrodynamical simulations show that a spherical mass ejection of $1\text{--}2\ M_{\odot}$ over a period of 20 years into a massive pre-existing torus of gas and dust can result in the present-day geometry^{6,7}. The gas ejected in 1843 shows a strong overabundance of nitrogen and depletion of carbon and oxygen, compared to a solar-composition gas^{19,20}. This abundance pattern is characteristic for gas in which nuclear burning of hydrogen via the CNO cycle has taken place. The star which ejected the mass must have these layers exposed at its surface, requiring the prior removal of the order of $10\text{--}30\ M_{\odot}$ of normal-composition material from the envelope of the star before the 1843 eruption; this removal could have occurred either in an extended period of mass loss in a stellar wind, by similar

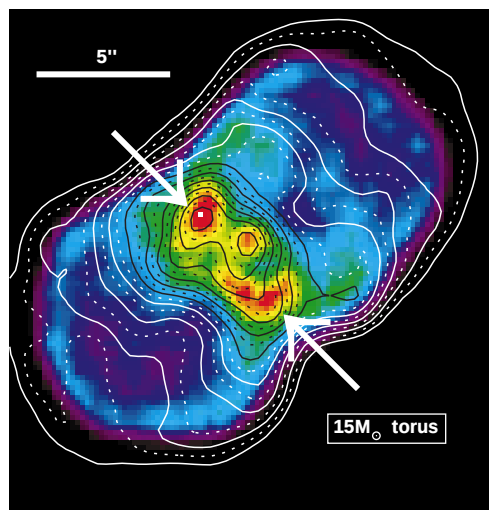


Figure 2 False-colour image of $\eta\ \text{Car}$ at $17\ \mu\text{m}$ wavelength. This deconvolved Thermal Infrared Multi-Mode Instrument (TIMMI) image ($\Delta\lambda = 1.5\ \mu\text{m}$) was observed in standard chopping/nodding mode. The detector used was a 64×64 Ga:Si array. The $17\text{--}\mu\text{m}$ image has a spatial resolution of $1.3''$ (full-width at half-maximum). The core is clearly structured and has three brightness peaks. Overplotted are contours from the deconvolved TIMMI image at $10\ \mu\text{m}$ ($\Delta\lambda = 5.2\ \mu\text{m}$), increasing linearly in brightness. The central peak is the power-law component, and is much brighter in the $10\text{--}\mu\text{m}$ image than in the $17\text{--}\mu\text{m}$ image. The minor axis (the 'waist') contains most of the cooler ($\sim 110\ \text{K}$) material. Comparison of the images also indicates that the surface brightness of the outer edge of the nebula diminishes towards the longer wavelengths, while the size of the nebula at $10\ \mu\text{m}$ (indeed at wavelengths at least to $12.8\ \mu\text{m}$; ref. 15) is spatially coincident with the Hubble Space Telescope images^{24,25}. The apparent difference in size revealed here at $17\ \mu\text{m}$ may be due to the presence of warm silicate grains (strongly contributing to the $10\text{--}\mu\text{m}$ silicate emission band), located exterior to the limb-brightened material. The limb-brightened material containing $15\ M_{\odot}$ is indicated by the arrows in the $17\text{--}\mu\text{m}$ image. The continuum dust emission in the layer between the limb-brightened material and the edge of the optical nebula is too weak to be detected in the $17\text{--}\mu\text{m}$ image.

events to those that occurred in 1843, or by interaction between the two stars in the system.

We suggest that the massive torus could be the normal-composition envelope of the star. If this is the case, then the expansion velocity of the torus suggests that the star lost its envelope some 1,000–2,000 years ago. The highly flattened geometry then points to interaction between the two stars in the system as a plausible mechanism for the removal of the envelope. We note that the age of the torus is a lower limit; it is possible that the present-day stellar wind of the star has radially accelerated the material in the torus. An older torus may have been formed by catastrophic mass loss if a common stellar envelope was created when one or both stars had convective envelopes in a previous (red-supergiant-like) evolutionary phase. If the stripped-envelope model is correct, then the gas in the torus should not be severely depleted of carbon and oxygen, as the lobes are. The origin of the 1843 ejection remains unclear, but the classical idea is that it may be related to the proximity of the star to the Eddington limit, possibly helped by rapid rotation³. At this limit, the force due to radiation pressure (possibly aided by centrifugal force⁸) exerted on the outer layers of the star equals the force on these layers due to gravity, which is expected to result in severe instabilities as evidenced by violent mass ejection. The sudden, binary-induced removal of around 15 $M_{\odot}$ from the stellar envelope drastically increases the ratio of total luminosity to stellar mass. This would quickly (on modified Kelvin–Helmholtz, or gravitational contraction, timescales of $\sim 10^3$ years) put the star close to the Eddington limit, perhaps beyond it. Massive stars allowed to evolve in solitude would reach this limit gradually, perhaps avoiding the type of eruption seen in η Car in the nineteenth century. $\square$

Received 23 June; accepted 4 October 1999.

1. Damineli, A. The 5.52 year cycle of Eta Carinae. *Astrophys. J.* **460**, L49–L52 (1996).
2. Davidson, K. Is Eta Carinae a long-period binary? *New Astron.* **2**, 387–396 (1997).
3. Davidson, K. & Humphreys, R. M. Eta Carinae and its environment. *Annu. Rev. Astron. Astrophys.* **35**, 1–32 (1997).
4. Nota, A. & Lamers, H. (eds) *Luminous Blue Variables: Massive Stars in Transition* (ASP Conf. Ser. 120, Astronomical Society of the Pacific, San Francisco, 1997).
5. Morse, J. A., Humphreys, R. M. & Damineli, A. (eds) *Eta Carinae at the Millennium* (ASP Conf. Ser. 179, Astronomical Society of the Pacific, in the press).
6. Frank, A., Balick, B. & Davidson, K. The Homunculus of Eta Carinae: an interacting stellar winds paradigm. *Astrophys. J.* **441**, L77–L80 (1995).
7. Dwarkadas, V. V. & Balick, B. On the formation of the Homunculus Nebula around η Carinae. *Astron. J.* **116**, 829–839 (1998).
8. Langer, N., García-Segura, G. & Mac Low, M.-M. Giant outbursts of luminous blue variables and the formation of the Homunculus Nebula around η Carinae. *Astrophys. J.* **520**, L49–L53 (1999).
9. Aitken, D. K., Smith, C. H., Moore, T. J. T. & Roche, F. Mid-infrared studies of η Carinae—II. Polarimetric imaging at 12.5 μ m and the magnetic field structure. *Mon. Not. R. Astron. Soc.* **273**, 359–366 (1995).
10. Duschl, W. J., Hofmann, K.-H., Rigaut, F. & Weigelt, G. Morphology and kinematics of Eta Carinae. *Rev. Mex. Astron. Astrof.* **27**, 1–22 (1995).
11. Falcke, H., Davidson, K., Hofmann, K.-H. & Weigelt, G. Speckle-masking imaging polarimetry of η Carinae: evidence for an equatorial disk. *Astron. Astrophys.* **306**, L17–L20 (1996).
12. Rigaut, F. & Gehring, G. The inner core of η Carinae in the near infrared. *Rev. Mex. Astron. Astrof.* **2**, 27–35 (1995).
13. Zethson, T. et al. Strange velocities in the equatorial ejecta of Eta Carinae. *Astron. Astrophys.* **344**, 211–220 (1999).
14. Cox, P. et al. Millimeter emission of Eta Carinae and its surroundings. *Astron. Astrophys.* **297**, 168–174 (1995).
15. Smith, N., Gehrz, R. D. & Krautter, J. The infrared morphology of η Carinae. *Astron. J.* **116**, 1332–1345 (1998).
16. Smith, N., Guzik, J. A., Lamers, H. J. G. L. M., Cassinelli, J. P. & Humphreys, R. M. in *Variable and Non-spherical Stellar Winds in Luminous Hot Stars* (ed. Wolf, B. et al.) 249–250 (IAU Symp. 169, Springer, Heidelberg, 1999).
17. Smith, C. H. et al. Mid-infrared studies of η Carinae—I. Subarcsecond imaging at 12.5 and 17 μ m. *Mon. Not. R. Astron. Soc.* **273**, 354–358 (1995).
18. Kessler, M. et al. The Infrared Space Observatory (ISO) mission. *Astron. Astrophys.* **315**, L27–L31 (1996).
19. Davidson, K., Dufour, R. J., Wallborn, N. R. & Gull, T. R. Ultraviolet and visual wavelength spectroscopy of the gas around Eta Carinae. *Astrophys. J.* **305**, 867–879 (1986).
20. Viotti, R., Rossi, L., Cassatella, A., Altamore, A. & Baratta, G. B. The ultraviolet spectrum of Eta Carinae. *Astrophys. J. Suppl.* **71**, 983–1009 (1989).
21. de Graauw, Th. et al. Observing with the ISO short-wavelength spectrometer. *Astron. Astrophys.* **315**, L49–L54 (1996).
22. Clegg, P. E. et al. The ISO long-wavelength spectrometer. *Astron. Astrophys.* **315**, L38–L42 (1996).
23. Harvey, P. M., Hoffmann, W. F. & Campbell, M. F. High angular resolution observations of η Carinae

from 35–175 microns. *Astron. Astrophys.* **70**, 165–168 (1978).

24. Humphreys, R. M. & Davidson, K. The luminous blue variables: astrophysical geysers. *Publ. Astron. Soc. Pacif.* **106**, 1025–1051 (1994).

25. Morse, J. et al. Hubble Space Telescope wide field planetary camera 2 observations of Eta Carinae. *Astron. J.* **116**, 2443–2461 (1998).

Acknowledgements

We appreciate the use of the composite HST images constructed by J. Morse. We also thank E. Lellouch for assistance with calibrations pertaining to Uranus, I. Yamamura for discussions, and S. F. Moseley for access to the recent unpublished KAO photometry of η Car. This work was supported in part by a Pioneer grant to L.B.F.M.W. and a NWO Spinoza grant to E.P.J. van den Heuvel. This work is partly based on observations with ISO, an ESA project with instruments funded by ESA member states and with the participation of ISAS and NASA. It also involves observations obtained at the European Southern Observatory, La Silla, Chile.

Correspondence and requests for materials should be addressed to P.W.M. (e-mail: p.morris@astro.uva.nl).

Observation of spin and charge collective modes in one-dimensional metallic chains

P. Segovia, D. Purdie, M. Hengsberger & Y. Baer

Institut de Physique, Université de Neuchâtel, CH-2000 Neuchâtel, Switzerland

The many-body theory of interacting electrons in solids establishes the existence of elementary excitations, named quasiparticles, which show a one-to-one correspondence with non-interacting electrons. But this so-called Fermi liquid approach breaks down spectacularly in one-dimensional metals¹. In this situation, which is described by the Luttinger liquid formalism, the quasiparticles are replaced by distinct collective excitations involving spin and charge, called spinons and holons, respectively². This approach predicts power-law behaviour for the various properties of one-dimensional metals which is experimentally testable using a wide variety of methods, such as transport measurements^{3,4} and optical conductivity measurements⁵. Photoemission, on the other hand, provides a means by which the spin and charge excitations can be observed directly. Previous photoemission studies of quasi-one-dimensional metals have essentially revealed only the absence of any discontinuity of the spectral function at the Fermi energy⁶, consistent with theoretical expectations. Recently, signatures of the existence of spin-charge separation have been inferred from line-shape analyses in a metal with different bands⁷ and in an insulator⁸. Here we present photoemission data from a genuine one-dimensional metal constructed on an insulating substrate. The spectra contain structures indicative of the excitation of spin and charge collective modes.

The use of photoemission-based techniques to search for Luttinger-liquid behaviour has been largely confined to anisotropic solids which exhibit a one-dimensional metallic character. In order to avoid uncertainties associated with crossover to higher-dimensional regimes in such samples, attempts have been made to explore one-dimensional structures existing on surfaces^{9,10}. Recently, chains of Au atoms on Si(111) have received some attention. Vicinal substrates have been used to inhibit multiple domain growth, but up until now the samples have not been monophase, containing regions of bunched steps also present with the desired structure¹¹. Furthermore, the relatively poor energy resolution of previous studies impedes a detailed examination of the spectra in the critical near-Fermi-level region^{12,13}. Here we examine chains of Au atoms on

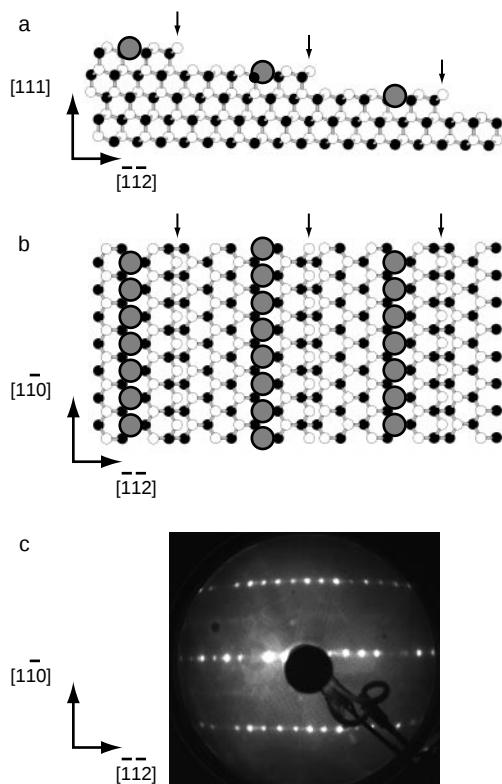


Figure 1 Surface crystallography of the system. Side (a) and plan (b) views of Si(557). In the plan view, only the two top layers of Si atoms (white and black circles) are shown. The surface displays (111) terraces separated by regular steps, marked with arrows. Au atoms are represented in the figure (large grey circles) in a sites consistent with coverage, symmetry and previous scanning tunnelling microscopy work¹⁶. c, The low-energy electron diffraction pattern of the surface. This represents an image of the reciprocal lattice of the surface.

Si(111), making two crucial advances. First, we are able to measure properties at low temperature (~ 10 K) and with an appropriately high instrumental energy resolution (5 meV)¹⁴. Second, for the first time, to our knowledge, photoemission is used to examine a uniform monophase sample with a regular step structure. The Si sample substrate (p-type B doped, $\rho_{300\text{ K}} = 10\ \Omega\text{ cm}$) is cut so that the macroscopic surface makes an angle of 9.45° with the (111) planes of Si, the misorientation critically chosen in the $[\bar{1}\bar{1}2]$ direction. A precise Au coverage of 0.2 monolayers (with respect to the surface atom concentration of Si(111)1 $\times$ 1) is evaporated onto the substrate, previously cleaned by flashing to $1,400^\circ\text{C}$. The sample is then annealed to 950°C to obtain a well ordered surface, denoted Si(111)5 $\times$ 1Au. Coverage is checked by X-ray photoelectron and Auger spectroscopy, and surface order and symmetry by low energy electron diffraction (LEED). The regularly stepped superstructure imposed by the offcut and stabilized by the presence of Au on the surface¹⁵ consists of terraces which are essentially five substrate lattice spacings wide (Fig. 1a and b). A LEED image of the surface lattice is shown in Figure 1c, in which the regular superstructure along the miscut direction is clearly revealed by the closely spaced diffraction spots. Although the detailed surface crystallography of our system is not known, the Au coverage allows us to deduce that one adsorbate atom is present in every surface unit cell. Thus a single adsorbate-induced band originating from the Au 6s level is expected to be populated in the ground state. The symmetry observed by LEED, which does not alter between 300 K and 60 K, ensures that the Au atoms on a terrace are linearly aligned, separated from each other by $3.83\ \text{\AA}$. The inter-chain spacing is equal to the terrace width and is of the order of $20\ \text{\AA}$, ensuring a

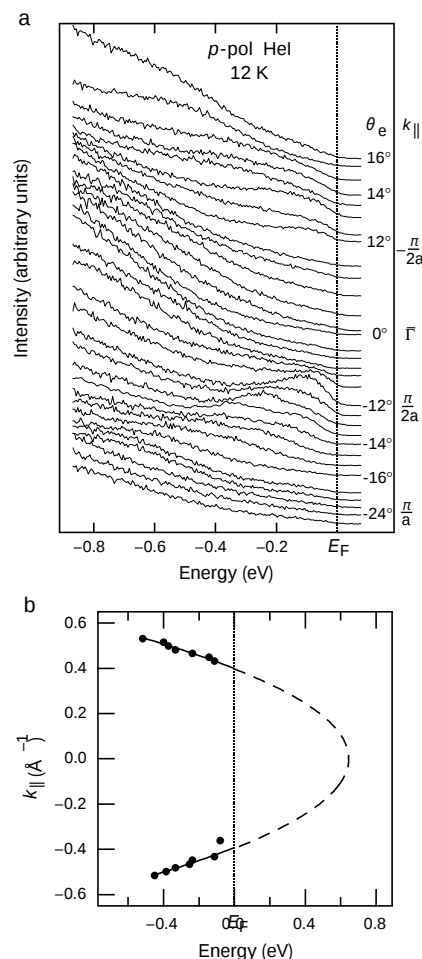


Figure 2 Dispersion of the Au-induced state parallel to the chains. a, Angle-resolved photoemission spectra of Si(111)5 $\times$ 1Au, with the surface wave vector ($k_{||}$) along the chain direction $[1\bar{1}0]$. The polar emission angle θ_e is noted with respect to the surface normal. The angular steps between spectra for which the value of θ_e is marked on the figure are equal. $k_{||}$ values are given for emission at E_F , with a being the inter-Au distance within a chain, $3.83\ \text{\AA}$. The electric vector of the radiation is contained in the emission plane (p -pol). In order to compensate for effects due to the polarization of the exciting radiation, negative angle spectra are multiplied by a factor of 0.75. An excitation feature is clearly seen dispersing towards zero energy as normal emission is approached. b, The energy and momentum positions of the main structure plotted and fitted to a parabola showing the characteristic dispersion of a metallic band.

complete lack of cross-chain interaction and a strictly one-dimensional structure. This picture of the surface is consistent with a model based on scanning tunnelling microscopy data¹⁶.

For our measurement, linearly polarized monochromatic ultraviolet photons ($\hbar\omega = 21.2\text{ eV}$) are incident on the sample. The photoemitted electrons are collected over a small angular range and analysed in energy. As the Au chains are located at the surface of our sample, the wave vectors of the electronic states linked to the chains lie in the surface plane. This momentum, $k_{||}$, is strictly conserved in the photoemission measurement. It is proportional to $\sqrt{E_k} \sin \theta_e$, where E_k is the kinetic energy of the emitted electron and θ_e is the angle between the emission direction and the surface normal¹⁷. The measurements are performed at 12 K in order to minimize the thermal broadening. The low temperature reduction in recombination rate of the radiation-induced charge creates a non-equilibrium surface photovoltage (SPV) tending to compensate for the band bending in the semi-conducting substrate¹⁸. Figure 2a shows stacked photoemission spectra from Si(111)5 $\times$ 1Au, corrected for any SPV shifts by comparison with room temperature data. These spectra are

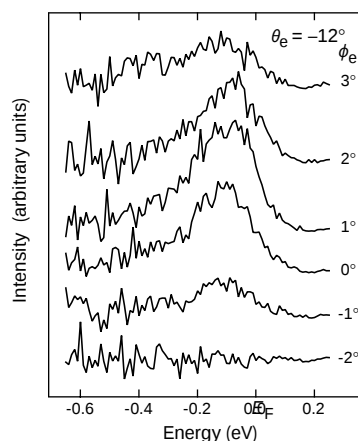


Figure 3 Angle-resolved photoemission spectra with varying surface wave vector $k_{\parallel}$ perpendicular to the chains. The polar angle θ_e in the emission plane is fixed at -12° and the angle ϕ_e in the perpendicular plane is varied. In order to remove any visual distortions coming from the featureless background in the raw spectra, this has been subtracted. The lack of dispersion in this direction confirms the one-dimensional nature of the system under study.

recorded with electron emission in the plane spanned by $[1\ 1\ 1]$ and $[1\ \bar{1}\ 0]$, and $k_{\parallel}$ is thus probed along the chain direction by rotating the sample. A dispersing state due to the Au/Si reconstruction can clearly be observed moving towards E_F as normal emission ($k_{\parallel} = 0$) is approached. The peak disappears in the range $-10^\circ \leq \theta_e \leq +10^\circ$, but the metallic character of this band can be inferred from its dispersion and occupation (Fig. 2b). The form of the photoemission spectra and, in particular, the dispersion behaviour displayed in Fig. 2 are unchanged between 300 K and 12 K. Specifically, no indication is observed suggestive of charge density wave (CDW) formation or any other temperature-induced phase transition in the range of study. This is consistent with the absence of a Peierls transition at finite temperature in an ideal one-dimensional system. The width of the cut-off at E_F in a photoemission spectrum from a metallic system exhibiting Fermi liquid behaviour is of the order of $k_B T$, where k_B is Boltzmann's constant, and T is temperature. Although, as argued above, the spectra presented in Fig. 2 demonstrate the presence of a metallic band, the measured high energy cut-offs are an order of magnitude larger than $k_B T$ at the temperature of measurement.

Angle-resolved photoemission spectra that were recorded while changing the k -vector perpendicular to the chain direction reveal no dispersion for the chain state (Fig. 3); this confirms that we have fabricated a one-dimensional metallic system which displays non-Fermi liquid spectral behaviour. The experimental facts indicate that our system is a realization of a Luttinger liquid, one of whose properties is vanishing spectral intensity at E_F (ref. 2).

The spectra of Fig. 2 showing the most pronounced structure are collected in Fig. 4. These spectra are recorded in the near k_F region, which is precisely where the Luttinger model can be expected to be valid. The relative intensity of the dispersing state is weak since it originates from only 20% of the total number of surface atoms. The relevant signals have been isolated by subtracting the featureless spectrum recorded at $\theta_e = -20^\circ$. In the region where the band crosses k_F ($\theta_e \sim -10^\circ$ to -12°), the spectra display a peak centred at ~ -100 meV. The intensity depletion replacing the metallic Fermi edge in these spectra is a feature systematically observed in all photoemission studies of one-dimensional metallic systems, and the spectral form raises a puzzling problem. At k_F , theory predicts a power law form for the momentum resolved spectral function, $\sim (E - E_F)^{\alpha-1}$ (refs 19, 20). Agreement of this with the experimental spectra can only be obtained for values of α largely exceeding one. This requires a strong long-range electron-electron interaction^{19,20},

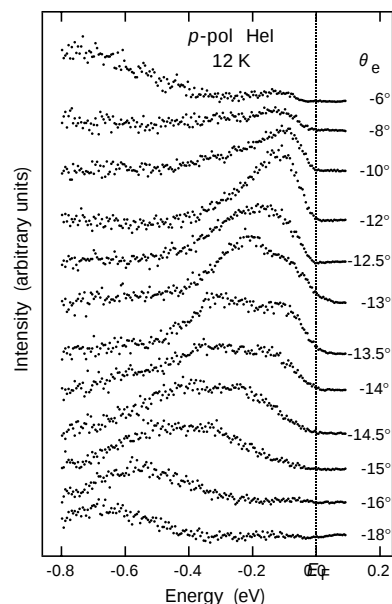


Figure 4 Detailed angle-resolved photoemission spectra of Si(111)5 $\times$ 1Au for θ_e between -6° and -18° . To emphasize the structures of interest, a featureless background has been removed as explained in the text. The double-peaked structure observed in the range $\theta_e = -13^\circ$ to -15° is assumed to originate from the excitation of spin and charge collective modes. The spinon lies closest to E_F , and the holon at deeper energy.

a condition unlikely to be realized in all one-dimensional systems studied up to now. In experiment, defects and impurities in one-dimensional samples set limits to the chain length. Calculations have shown that such boundary conditions markedly increase this exponent α for small energies and momenta²¹. This mechanism alone does not seem sufficient to fully account for the spectral shape near E_F , but other experimental conditions can also be invoked. The detailed spectral form, in particular the edge shape, depends critically upon the sample preparation, the best data being obtained very close to the threshold concentration of Au which induces the cancellation of SPV. The macroscopic surface probed by photoemission (~ 1 mm²) cannot be strictly homogeneous, and it is anticipated that the spectra are formed by superimposed signals which originate from regions at slightly different SPV, contributing a smearing of the edge shape. For these reasons it is not possible to extract a meaningful value of α from the high energy cut-off at k_F of our spectra. The recent suggestion²² that a pseudo-gap could be falsely observed in samples with poorly conducting substrates due to ohmic losses induced by the photoemission process is unlikely to be important in our measurements. This mechanism requires low-energy losses which do not exist in our insulating substrate. In addition, the structures originating from the one-dimensional metallic chains show no sizeable asymmetric loss of intensity when the temperature increases from 12 K to 300 K. For k -vectors up to k_F ($\theta_e = -6^\circ$, -8° and -10°), a peak emerges as E_F is approached. For further increasing k the peak initially shows an anomalous broadening ($\theta_e = -12.5^\circ$), and then a two-peak structure appears ($\theta_e = -13^\circ$) with a dominant weight at low energy. In the subsequent spectra ($\theta_e = -13.5^\circ$ to -15°) a very unusual line shape is observed. From both sides, the high and low energy flanks are followed by maxima with a plateau in between. The relative intensity variation between the two features, as well as their junction when reaching E_F , seems incompatible with a spin-split origin²³. The aspect and evolution with angle of the spectra are strongly reminiscent of the predicted behaviour for the excitation of spin and charge modes^{6,20,24}. The experimental structures are strongly attenuated when compared to calculations, but their survival is not incompatible with the large value of α inferred from the

threshold behaviour. Approximate theoretical considerations (S. Eggert, personal communication) suggest that for finite values of $(k-k_F)$, experimentally imposed boundary conditions do not destroy the spin and charge excitations and leave their positions unmodified. However, their singular behaviour is expected to be markedly damped, as observed in Fig. 4. This softening of the singularities in the experimental spectra, when compared to the Luttinger-liquid spectral functions calculated with periodic boundary conditions^{6,20,24}, must be attributable to imperfections in the chains and to inhomogeneities of the surface, and not to our instrumental resolution of 5 meV. The theoretical predictions of the Luttinger model are limited to small energies around E_F . If we interpret the two peaks as holon and spinon excitations in the narrow range ($\theta_c \approx -13^\circ$ to -14°), then from their approximate linear behaviour we can estimate a value between 2 and 3 for their velocity ratio. This corresponds to a range of α value compatible with the observation of distinct peaks²⁰. The persistence of spin and charge excitations relatively far away from the region of linear band dispersion around k_F has been confirmed by calculations²⁵. In the two last spectra ($\theta_c = -16^\circ$ and -18°) where the Brillouin zone boundary ($\theta_c \approx -24^\circ$) is approached, the two structures join in an asymmetric peak with its maximum in line with the holon dispersion. $\square$

Received 4 June; accepted 7 October 1999.

1. Luttinger, J. M. An exactly soluble model of a many-fermion system. *J. Math. Phys.* **4**, 1154–1162 (1963).
2. Voit, J. One-dimensional Fermi liquids. *Rep. Prog. Phys.* **18**, 997–1116 (1995).
3. Grayson, M., Tsui, D. C., Pfeiffer, L. N., West, K. W. & Chang, A. M. Continuum of chiral Luttinger liquids at the fractional quantum hall edge. *Phys. Rev. Lett.* **80**, 1062–1065 (1998).
4. Bockrath, M. et al. Luttinger-liquid behaviour in carbon nanotubes. *Nature* **397**, 598–601 (1999).
5. Schwartz, A. et al. On-chain electrodynamics of metallic (TMTSF)₂X salts: observation of Tomonaga-Luttinger liquid response. *Phys. Rev. B* **58**, 1261–1271 (1998).
6. Griener, M. & Voit, J. in *Electron Spectroscopies Applied to Low-dimensional Materials* (eds Stanberg, H. & Hughes, H.) (Kluwer, Dordrecht, in the press).
7. Denlinger, J. D. et al. Non-Fermi-liquid single particle line shape of the quasi-one-dimensional non-CDW metal Li_{0.9}Mo_{0.17}: comparison to the Luttinger liquid. *Phys. Rev. Lett.* **82**, 2540–2543 (1999).
8. Kim, C. et al. Observation of spin-charge separation in one-dimensional SrCuO₂. *Phys. Rev. Lett.* **77**, 4054–4057 (1996).
9. Bertel, E. & Lehmann, L. Electronic structure of self-organised lateral superlattices on a metal surface: O/Cu(110). *Phys. Rev. Lett.* **80**, 1497–1500 (1998).
10. Hill, I. G. & McLean, A. B. Metallicity of In chains on Si(111). *Phys. Rev. B* **56**, 15725–15728 (1997).
11. O'Mahony, J. D., McGilp, J. F., Flippe, C. F. J., Weightman, P. & Leible, F. M. The nucleation and evolution of the 5 × 2-Au structure on Si(111). *Phys. Rev. B* **49**, 2527–2535 (1994).
12. Collins, I. R. et al. Angle-resolved photoemission from an unusual quasi-one-dimensional metallic system: a single domain Au-induced 5 × 2 reconstruction of Si(111). *Surf. Sci.* **325**, 45–49 (1995).
13. Hill, I. G. & McLean, A. B. Role of adatoms in the Si(111)-Au(5 × 2) quasi-one-dimensional system. *Phys. Rev. B* **55**, 15664–15668 (1997).
14. Garnier, M., Purdie, D., Breuer, K., Hengsberger, M. & Baer, Y. Momentum resolved photoemission of the Kondo peak in an ordered Ce-containing alloy. *Phys. Rev. B* **56**, 11399–11402 (1997).
15. Jalochofski, M., Strozak, M. & Zdyb, R. Gold-induced ordering on vicinal Si(111). *Surf. Sci.* **375**, 203–209 (1997).
16. Shibata, M., Sumita, I. & Nakajima, M. Structure and width of Au-adsorbed narrow Si(111) terraces. *Phys. Rev. B* **57**, 1626–1630 (1998).
17. Kevan, S. D. (ed.) *Angle-Resolved Photoemission: Theory and Current Applications* (Elsevier, Amsterdam, 1992).
18. Demuth, J. E., Thompson, W. J., DiNardo, N. J. & Imbühl, R. Photoemission-based photovoltage probe of semiconductor surface and interface electronic structure. *Phys. Rev. Lett.* **56**, 1408–1411 (1986).
19. Voit, J. in *The Physics and Mathematical Physics of the Hubbard Model* (eds Baeriswyl, D., Campbell, D. K., Carmelo, J. M. P., Guinea, F. & Louis, E.) 263–272 (Plenum, New York, 1995).
20. Voit, J. Charge-spin separation and the spectral properties of Luttinger liquids. *J. Phys. C* **5**, 8305–8336 (1993).
21. Eggert, S., Johannesson, H. & Mattsson, A. Boundary effects on spectral properties of interacting electrons in one dimension. *Phys. Rev. Lett.* **76**, 1505–1508 (1996).
22. Joynt, R. Pseudogaps and extrinsic losses in photoemission experiments on poorly conducting solids. *Science* **284**, 777–779 (1999).
23. LaShell, S., McDougall, B. A. & Jensen, E. Spin splitting of an Au(111) surface state band observed with angle resolved photoelectron spectroscopy. *Phys. Rev. Lett.* **77**, 3419–3422 (1996).
24. Meden, V. & Schönhammer, K. Spectral functions for the Tomonaga-Luttinger model. *Phys. Rev. B* **46**, 15753–15760 (1992).
25. Zacher, M. G., Arrigoni, E., Hanke, W. & Schrieffer, J. R. Systematic numerical study of spin-charge separation in one dimension. *Phys. Rev. B* **57**, 6370–6375 (1998).

Acknowledgements

We thank H. Beck and R. Frésard for discussions. This work was supported by the Swiss National Science Foundation.

Correspondence and requests for materials should be addressed to P.S.
(e-mail: Segovia.Pilar@iph.unine.ch).

Resonant intermolecular transfer of vibrational energy in liquid water

Sander Woutersen & Huib J. Bakker

FOM-Institute for Atomic and Molecular Physics, Kruislaan 407,
1098 SJ Amsterdam, The Netherlands

Many biological, chemical and physical processes involve the transfer of energy. In the case of electronic excitations, transfer between molecules is rapid, whereas for vibrations in the condensed phase, resonant energy transfer is an unlikely process because the typical timescale of vibrational relaxation (a few picoseconds) is much shorter than that of resonant intermolecular vibrational energy transfer^{1,2}. For the OH-stretch vibration in liquid water, which is of particular importance due to its coupling to the hydrogen bond, extensive investigations have shown that vibrational relaxation takes place with a time constant of 740 ± 25 femtoseconds (ref. 7). So for resonant intermolecular energy transfer to occur in liquid water, the interaction between the OH-stretch modes of different water molecules needs to be extremely strong. Here we report time-resolved pump-probe laser spectroscopy measurements that reveal the occurrence of fast resonant intermolecular transfer of OH-stretch excitations over many water molecules before the excitation energy is dissipated. We find that the transfer process is mediated by dipole-dipole interactions (the Förster transfer mechanism⁹) and additional mechanisms that are possibly based on intermolecular anharmonic interactions involving hydrogen bonds. Our findings suggest that liquid water may play an important role in transporting vibrational energy between OH groups located on either different biomolecules or along extended biological structures. OH groups in a hydrophobic environment should accordingly be able to remain in a vibrationally excited state longer than OH groups in a hydrophilic environment.

To investigate the occurrence of resonant intermolecular energy transfer in liquid water, we have used polarization-resolved vibrational pump-probe spectroscopy to study both liquid H₂O and solutions of HDO in D₂O with different HDO concentrations (where D is deuterium). Our experimental system^{7,10} employs optical parametric generation and amplification to obtain two independently tunable intense mid-infrared pulses with a duration of approximately 200 fs. One of these pulses has an energy of 20 μJ and is used as a pump pulse, the other has an energy of less than 1 μJ and is used as a probing pulse. The pump pulse is tuned to the $\nu_{OH} = 0 \rightarrow 1$ transition frequency at $3,400 \text{ cm}^{-1}$ ($\lambda = 2.94 \text{ μm}$) to excite a small fraction of the water molecules to the first excited state of the OH-stretch mode. This leads to an absorption increase $\Delta\alpha$ at the $\nu_{OH} = 1 \rightarrow 2$ transition frequency, which is the absorption frequency of the vibrationally excited OH-oscillators. This absorption increase is monitored by the probing pulse (centre frequency $3,000 \text{ cm}^{-1}$, $\lambda = 3.33 \text{ μm}$) as a function of the delay τ between pump and probe. The $\nu_{OH} = 0 \rightarrow 1$ and $\nu_{OH} = 1 \rightarrow 2$ transition frequencies are well separated owing to the anharmonicity of the OH-potential⁵.

The samples consist of a thin layer (1–25 μm) of either pure H₂O (HPLC grade) or HDO dissolved in D₂O, and are kept between two CaF₂ windows. The HDO:D₂O samples are prepared by mixing appropriate amounts of H₂O (HPLC grade) and D₂O

(>99.9 atom% D). All experiments are performed at room temperature (298 K). The samples are rapidly rotated during the experiments to avoid heating in the focus.

The pump pulse is linearly polarized, and OH-oscillators directed along the pump polarization are preferentially excited, producing an anisotropic orientational distribution of the excited OH-groups^{4,11}. As a result, the absorption increase $\Delta\alpha$ at the $\nu_{\text{OH}} = 1 \rightarrow 2$ frequency will initially be larger for light polarized parallel to the pump polarization ($\Delta\alpha_{\parallel}$) than for light polarized perpendicular to the pump polarization ($\Delta\alpha_{\perp}$). The delay dependence of $\Delta\alpha_{\parallel}$ and $\Delta\alpha_{\perp}$ is determined by both the population relaxation of the $\nu_{\text{OH}} = 1$ state, which causes a decrease of both $\Delta\alpha_{\parallel}$ and $\Delta\alpha_{\perp}$, and the depolarization of the OH-stretch excitation, which causes a decrease of the difference between $\Delta\alpha_{\parallel}$ and $\Delta\alpha_{\perp}$. The rotational anisotropy is defined as:

$$R(\tau) = \frac{\Delta\alpha_{\parallel}(\tau) - \Delta\alpha_{\perp}(\tau)}{\Delta\alpha_{\parallel}(\tau) + 2\Delta\alpha_{\perp}(\tau)} \quad (1)$$

It can be shown that $R(\tau)$ is not sensitive to the population relaxation, but only to the depolarization¹¹. The polarization of the probe pulse is rotated 45° with respect to that of the pump pulse, and the absorption changes parallel ($\Delta\alpha_{\parallel}$) and perpendicular ($\Delta\alpha_{\perp}$) to the pump polarization are detected simultaneously, which allows us to determine the rotational anisotropy as a function of delay on a single-shot basis¹².

Figure 1 shows the decay of the rotational anisotropy as a function of delay observed in pure H₂O and in solutions of HDO in D₂O with three different HDO concentrations. Clearly, the decay of the rotational anisotropy of the OH-stretch excitation becomes faster with increasing concentration of OH oscillators. This concentration dependence of the decay of the rotational anisotropy shows that resonant intermolecular energy transfer of the OH-stretch excitation occurs in liquid water. The rate of intermolecular energy transfer and thus of depolarization strongly increases with decreasing average distance between the OH groups. Hence, we find that the high concentration of OH-oscillators leads to strong intermolecular interactions that render the vibrational dynamics essentially different from that observed for strongly diluted HDO:D₂O (refs 4–8).

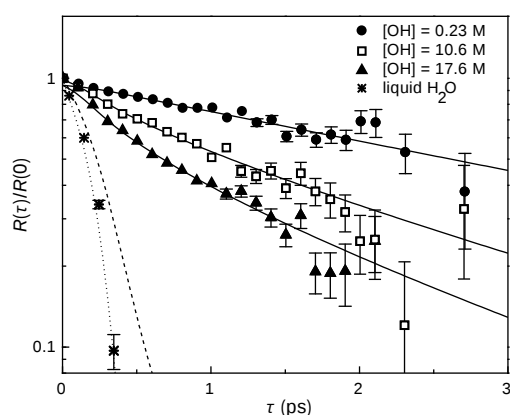


Figure 1 Rotational anisotropy $R(\tau)/R(0)$ as a function of delay. Data is normalized with respect to the value at zero delay, as observed in HDO:D₂O solutions with three different HDO concentrations, and in pure H₂O. The solid curves are obtained by least-squares fitting of equation (3) to all of the HDO:D₂O data, using $T_1 = 740$ fs (ref. 7) and the appropriate values for [OH] for each of the three data sets, and taking into account the pulse duration of approximately 200 fs. This least-squares fitting procedure resulted in best-fit parameter values of 4.0 ± 0.4 ps for τ_{or} and 2.10 ± 0.05 Å for the Förster radius r_0 . The dashed curve is calculated using these values and an OH concentration of 111 M. The dotted curve is calculated assuming an instantaneous decay of the rotational anisotropy.

In a quantitative analysis of the data, it should be realized that the decay of the rotational anisotropy is caused not only by intermolecular transfer of the OH-stretch excitation, but also by the orientational diffusion of the OH-groups⁴. Obviously, the latter contribution to the depolarization will not depend on the concentration. At very low HDO concentration, the intermolecular energy transfer becomes negligible, and the decay of the anisotropy results from orientational diffusion only. In a previous study on the orientational diffusion of water molecules⁴, we found that this process takes place on two distinct timescales, a slow timescale associated with the reorientation of strongly hydrogen-bonded molecules and a fast timescale for weakly hydrogen-bonded molecules. In these experiments, the frequency of the probe pulse was tuned to the $\nu_{\text{OH}} = 0 \rightarrow 1$ transition, and consequently the orientational dynamics of both molecules in the vibrational ground ($\nu_{\text{OH}} = 0$) state and molecules in the first excited ($\nu_{\text{OH}} = 1$) state were probed. In the present experiments, the probe frequency is tuned to the $\nu_{\text{OH}} = 1 \rightarrow 2$ transition, and only the excited ($\nu_{\text{OH}} = 1$) molecules are probed. These excited molecules are more strongly hydrogen-bonded than the molecules in the $\nu_{\text{OH}} = 0$ state^{13,14}, so that the contribution of the fast component associated with the weakly hydrogen-bonded molecules becomes negligible. As a result only the slow timescale is observed for the orientational motion, so that at low HDO concentration the decay of the anisotropy can be well described by a single exponential.

If the intermolecular energy transfer is described by a dipole–dipole (Förster-type) interaction⁹, the rate of energy transfer k between two OH-oscillators is given by

$$k = T_1^{-1}(r_0/r)^6, \quad (2)$$

where T_1 is the lifetime of the excited state, r the distance between the oscillators, and r_0 a constant usually referred to as the Förster radius. The Förster radius r_0 is an important material parameter that characterizes the intermolecular energy transfer, and that as yet has not been determined for water. From equation (2) we see that if r_0 is smaller than the average distance between the OH groups, no significant energy transfer will occur during the lifetime T_1 of the excitation. If r_0 is comparable to or larger than the average distance between the OH groups, the energy transfer will become significant for the dynamical behaviour. Using equation (2) and incorporating the orientational diffusion, the following equation for the decay of the rotational anisotropy R with increasing delay τ can be derived¹⁵:

$$\frac{R(\tau)}{R(0)} = \exp\left(\sqrt{\frac{\tau}{\tau_{\text{or}}}} - \frac{4\pi^{3/2}}{3}[\text{OH}] \frac{r_0^6 \tau}{T_1}\right) \quad (3)$$

where [OH] is the concentration of molecules containing an OH group. In the HDO:D₂O solutions, [OH] is nearly equal to the HDO concentration, but also contains a small contribution due to H₂O molecules. We found that all delay scans recorded in HDO:D₂O solution can be well described by equation (3). By least-squares fitting of this equation to the HDO:D₂O data, shown as solid curves in Fig. 1, we obtained $\tau_{\text{or}} = 4.0 \pm 0.4$ ps, and a Förster radius $r_0 = 2.10 \pm 0.05$ Å. This value of $\tau_{\text{or}} = 4.0 \pm 0.4$ ps represents a much more accurate determination of the time constant associated with slow reorientation than the value of 13 ps we reported previously⁴. We have significantly improved the signal-to-noise ratio of our data acquisition, which has enabled us to determine the time constant of the slow reorientation process far more accurately. The value of 2.10 Å that we find for r_0 is comparable to the intermolecular distance of ~ 2.8 Å in liquid water¹⁶, which implies that intermolecular energy transfer will be very fast in the pure liquid.

In Fig. 1 we have also plotted the decay of R predicted for liquid H₂O by the Förster mechanism (equation (3)). We used the r_0 and τ_{or} values obtained from the least-squares fit to the HDO:D₂O data and an OH concentration of 111 M (dashed curve), which is twice

the concentration of H₂O molecules in liquid H₂O, since each H₂O molecule has two OH-stretch modes (symmetric and asymmetric) and hence should be counted twice. The experimentally observed decay occurs faster than predicted by the Förster mechanism. In fact, we find that the H₂O data can be well described by assuming an instantaneous decay of the rotational anisotropy (dotted curve). This means that in liquid H₂O the intermolecular energy transfer takes place faster than our time resolution of approximately 100 fs. The rapid decay of *R* cannot be caused by intramolecular redistribution of the OH-stretch excitation between the symmetric and asymmetric OH-stretch modes of the H₂O molecule, since this can at most lead to a decay of the rotational anisotropy to a quarter of its initial value¹⁸. We note that there will be a slight difference in the orientational diffusion time constants τ_{or} of H₂O and HDO. It has been determined that the ratio of the time constants for orientational diffusion of H₂O and D₂O is approximately 0.8 at room temperature¹⁷. Since the moment of inertia of HDO has a value between those of H₂O and D₂O, the ratio of the time constants for HDO and H₂O will thus lie between 0.8 and 1. This small difference in τ_{or} for HDO and H₂O is not relevant in our experiments, since in H₂O the intermolecular energy transfer occurs on a much faster timescale than the orientational diffusion, which means that in H₂O the contribution of the latter to the decay of the rotational anisotropy is completely negligible.

The extremely high rate of resonant energy transfer in liquid H₂O is probably a consequence of the OH groups in liquid H₂O being so close to each other that not only dipole–dipole, but also dipole–quadrupole, quadrupole–quadrupole, and higher-order interactions will become important. In addition, anharmonic coupling between the OH-oscillators of different molecules may contribute to the energy transfer, possibly through the hydrogen bond connecting them. The Förster expression hinges on the fact that these other interactions decrease much more rapidly with increasing distance than the dipole–dipole interaction, and thus become less important. At the short distances occurring in H₂O this is of course not the case, and a more general treatment of the intermolecular coupling is necessary to accurately describe this liquid. □

Received 25 May; accepted 12 October 1999.

- Vlahoyannis, Y. P., Krueger, H., Knudsen, J. T. & Weitz, E. Vibration-vibration energy transfer processes in liquid xenon: a measurement of the rate constant for $\text{HCl}(\nu = 2) + \text{HCl}(\nu = 0) \rightarrow 2\text{HCl}(\nu = 1)$. *Chem. Phys. Lett.* **121**, 272–278 (1985).
- Apkarian, V. A., Wiedeman, L., Janiesch, W. & Weitz, E. Vibrational energy transfer and migration processes in matrix isolated CH₃F. *J. Chem. Phys.* **85**, 5593–5610 (1986).
- Hadži, D. & Bratos, S. in *The Hydrogen Bond* (eds Schuster, P., Zundel, G. & Sandorfy, C.) Vol. II, Ch. 12 (Elsevier, Amsterdam, 1976).
- Woutersen, S., Emmerichs, U. & Bakker, H. J. Femtosecond mid-infrared pump-probe spectroscopy of liquid water: evidence for a two-component structure. *Science* **278**, 658–660 (1997).
- Graener, H., Seifert, G. & Laubereau, A. New spectroscopy of water using tunable picosecond pulses in the infrared. *Phys. Rev. Lett.* **66**, 2092–2095 (1991).
- Laenen, R., Rauscher, C. & Laubereau, A. Dynamics of local substructures in water observed by ultrafast infrared hole burning. *Phys. Rev. Lett.* **80**, 2622–2625 (1998).
- Woutersen, S., Emmerichs, U., Nienhuys, H.-K. & Bakker, H. J. Anomalous temperature dependence of vibrational lifetimes in water and ice. *Phys. Rev. Lett.* **81**, 1106–1109 (1998).
- Gale, G. M. *et al.* Femtosecond dynamics of hydrogen bonds in liquid water: a real time study. *Phys. Rev. Lett.* **82**, 1068–1071 (1999).
- Förster, T. in *Modern Quantum Chemistry* Vol. III (ed. Sinanoglu, O.) 93–137 (Academic, New York, 1965).
- Emmerichs, U., Woutersen, S. & Bakker, H. J. Generation of intense femtosecond optical pulses around 3 μm with kHz rep-rate. *J. Opt. Soc. Am. B* **14**, 1480–1483 (1997).
- Graener, H., Seifert, G. & Laubereau, A. Direct observation of rotational relaxation times by time-resolved infrared spectroscopy. *Chem. Phys. Lett.* **172**, 435–439 (1990).
- Nienhuys, H.-K., Woutersen, S., van Santen, R. A. & Bakker, H. J. Mechanism for vibrational relaxation in water investigated by femtosecond infrared spectroscopy. *J. Chem. Phys.* **111**, 1494–1500 (1999).
- Stepanov, B. I. Interpretation of the regularities in the spectra of molecules forming the intermolecular hydrogen bond by the predissociation effect. *Nature* **157**, 808 (1946).
- Staib, A. & Hynes, J. T. Vibrational predissociation in hydrogen-bonded OH...O complexes via OH stretch–OO stretch energy transfer. *Chem. Phys. Lett.* **204**, 197–205 (1993).
- Eisenthal, K. B. Measurement of intermolecular energy transfer using picosecond light pulses. *Chem. Phys. Lett.* **6**, 155–157 (1970).
- Franks, F. (ed.) *Water, A Comprehensive Treatise* (Plenum, New York, 1972).
- Ronne, C., Åstrand, P.-O. & Keiding, S. THz spectroscopy of liquid H₂O and D₂O. *Phys. Rev. Lett.* **82**, 2888–2891 (1999).

- Szabo, A. Theory of fluorescence depolarization in macromolecules and membranes. *J. Chem. Phys.* **81**, 150–167 (1984).

Acknowledgements

We thank H. Schoenmaker for technical support and W. van der Zande and P. van der Meulen for critically reading the manuscript. This work is part of the research programme of the “Stichting voor Fundamenteel Onderzoek der Materie” (FOM), which is financially supported by the “Nederlandse organisatie voor Wetenschappelijke Onderzoek” (NWO).

Correspondence and requests for materials should be addressed to S.W.
(e-mail: s.woutersen@amolf.nl).

Cyclic transmembrane charge transport by pyrylium ions in a vesicle-based photocatalytic system

Rafail F. Khairutdinov & James K. Hurst

Department of Chemistry, Washington State University, Pullman, Washington 99164-4630, USA

Solar energy is being used for power generation, but also attracts increasing interest as a renewable energy source for the photocatalytic production of useful chemicals. Simple systems based on vesicles with transmembrane redox mediators have been used to transform photon energy into long-lived, membrane-separated photoredox products^{1–6}. However, these systems are not suitable for high-throughput applications because the transmembrane electron carriers are oxidized inside the vesicle into charged species that are no longer able to readily traverse the membrane bilayer. This leads to continuous trapping of these carriers during photolysis and, ultimately, to the termination of the redox reaction due to accumulation of the available carriers within the vesicle interior. Living cells circumvent this problem by using quinones to simultaneously transport electrons and protons, thus allowing the carrier to remain neutral in its reduced and oxidized states and so retain the ability to undergo transmembrane diffusion throughout the redox cycle. But the incorporation of quinones into artificial systems is not practical because of their susceptibility to oxidative degradation and slow transmembrane diffusion⁷. Here we describe an alternative mechanism for rapid electroneutral charge transport across vesicle membranes: we use pyrylium cations as the electron carrier, which undergo reversible ring-opening hydrolysis to form neutral diketones after deposition of the electron inside the vesicle. As the pyrylium cations are also the primary acceptors for the photoproduced electrons, our approach greatly simplifies the design of vesicle-based photocatalytic devices.

The complete system for pyrylium-ion-mediated transport of electrons and hydroxide ions is illustrated in Fig. 1. Ultracentrifugation studies⁴ have established that binding of ZnTPPS^{4–} to the anionic dihexadecyl phosphate (DHP) vesicles is negligible, whereas 92% of the Co(bpy)₃³⁺ and 94% of the 2,4,6-triphenylpyrylium ion, TPP⁺, were vesicle-bound under the experimental conditions. (Here ZnTPPS^{4–} is the [5,10,15,20-tetrakis(4-sulphonatophenyl)-porphinato]zinc(II) ion, and Co(bpy)₃³⁺ is the tris(2,2′-bipyridine)-Co(III) ion.) Binding of TPP⁺ was accompanied by a ~10-nm bathochromic shift in its lowest-energy electronic bands ($\lambda_{\text{max}}(\text{H}_2\text{O}) \sim 360 \text{ nm}$ and $\sim 410 \text{ nm}$), very similar to shifts induced by nonpolar solvents, indicating that the TPP⁺ ions are located in a relatively hydrophobic microenvironment within the membrane. The apparent equilibrium constant for hydrolysis of the pyrylium

ion to the uncharged pseudobase (TPPD, Fig. 1)^{8,9}, that is, $\text{TPP}^+ + \text{H}_2\text{O} = \text{TPPD} + \text{H}^+$, was evaluated spectrophotometrically in the presence of vesicles to be $K = [\text{TPPD}]/[\text{TPP}^+] > n20$ in 20 mM Tris/Cl, pH 8.0, and $K = 0.2$ in 10 mM Tris/Cl, 50 mM NaOAc, pH 5.0.

Continuous illumination of the complete system with wavelengths λ in the ZnTPPS^{4-} Soret (410 < λ < 430 nm) or Q bands ($\lambda > 545$ nm) gave rise to net reduction of the occluded $\text{Co}(\text{bpy})_3^{3+}$ to $\text{Co}(\text{bpy})_3^{2+}$ (Fig. 2). Photoreduction of $\text{Co}(\text{bpy})_3^{3+}$ was observed only when all of the reaction components were present. This reaction proceeded to completion when illuminated at $\lambda > 545$ nm. However, illumination in the near-ultraviolet region caused slow irreversible photobleaching of TPP^+ with an overall quantum yield of ~ 0.03 which competed with ZnTPPS^{4-} photoexcitation. Consequently, at low $[\text{TPP}^+]/[\text{Co}(\text{bpy})_3^{3+}]$ ratios, the extent of reaction under these illumination conditions became limited by TPP^+ photodecomposition. Control experiments established that no net topographic redistribution of ZnTPPS^{4-} , TPP^+ , or unreacted $\text{Co}(\text{bpy})_3^{3+}$ had occurred over the course of the reaction, although slow thermal leakage of the product $\text{Co}(\text{bpy})_3^{2+}$ ions from the vesicle interior was noted, consistent with earlier reports of similar behaviour for $\text{Ru}(\text{bpy})_3^{2+}$ ions in DHP vesicles¹⁰.

Electron transfer from excited ZnTPPS^{4-} (that is, $^3\text{ZnTPPS}^{4-}$) to TPP^+ is exothermic ($E^0(\text{ZnTPPS}^{3-}/^3\text{ZnTPPS}^{4-}) = -0.75$ V (NHE)¹¹, ($E^0(\text{TPP}^+/\text{TPP}^0) = -0.13$ V (NHE))¹². Both in aqueous solution and DHP vesicle suspensions, addition of TPP^+ caused the decay rate of $^3\text{ZnTPPS}^{4-}$ to increase, transient kinetic spectroscopy revealed simultaneous formation and subsequent decay of the ZnTPPS^{3-} π -cation radical¹³, indicating that oxidative quenching of the photoexcited triplet by pyrylium ions had occurred. No net changes were observed in the ground-state or transient absorption spectra or the intermediate decay rates after exposure to more than 100 laser flash excitations, indicating that the photoinitiated electron transfer reaction was fully reversible. Quenching rate constants, determined from the concentration dependence of the triplet lifetime, were $3.5 \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$ and

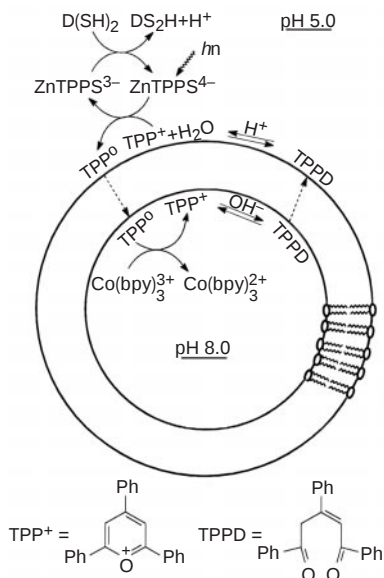


Figure 1 Reaction scheme for pyrylium-mediated electroneutral transmembrane charge transfer. An electron acceptor ($\text{Co}(\text{bpy})_3^{3+}$) in the inner aqueous phase of a dihexadecyl phosphate (DHP) small unilamellar vesicle is separated by a membrane from the other components. These other components, in the bulk aqueous phase, consist of an electron donor (dithiothreitol), photosensitizer (ZnTPPS^{4-}) and oxidative quencher/redox mediator (TPP^+). The internal and external aqueous solutions were usually 20 mM Tris/Cl, pH 8.0, and 10 mM Tris/Cl, 50 mM sodium acetate, pH 5.0, respectively.

$1.4 \times 10^8 \text{ M}^{-1} \text{ s}^{-1}$ in 20 mM Tris at pH 8.0 and in the DHP vesicular suspensions, respectively. The decrease in quenching constant was proportionate to the extent of TPP^+ binding to the vesicles, suggesting that the vesicle-bound TPP^+ ions were markedly less reactive toward $^3\text{ZnTPPS}^{4-}$ than were free TPP^+ ions.

Transmembrane diffusion of TPP^0 was monitored by measuring the rate of oxidation of externally added TPP^0 by $\text{Co}(\text{bpy})_3^{3+}$ ions located within the inner aqueous phase of the vesicles. For these experiments, both ZnTPPS^{4-} and the sacrificial electron donor, dithiothreitol, were not incorporated in the complete system. The reduction potential of singlet-excited TPP^+ —evaluated from $E^0(^1\text{TPP}^+/\text{TPP}^0) = E^0(\text{TPP}^+/\text{TPP}^0) + E_{00} = 2.67$ V, where $E_{00} = 2.80$ V is the energy of the first excited state of TPP^+ ¹⁴—is sufficiently large to allow oxidation of acetate ion ($E^0 = 2.1$ V)¹⁵ or organic impurities that might be present in the DHP. Laser pulse excitation at 355 nm produced a long-lived species with a visible absorption band centred at 550 nm, characteristic of the TPP^0 radical¹². The transient decay time decreased from $t_c^0 = 37$ ms in the absence of $\text{Co}(\text{bpy})_3^{3+}$ to $t_c = 3.4$ ms at $[\text{Co}(\text{bpy})_3^{3+}] \geq 5 \mu\text{M}$ (Fig. 3). Assuming a two-channel model for TPP^0 decay, we obtain an apparent transmembrane diffusion time (τ), given by $t_c \times t_c^0/(t_c^0 - t_c)$, of 3.7 ms. Assuming a bilayer width (l) of 4 nm, this value corresponds to a permeability coefficient (P) of $P = l/\tau \approx 1.1 \times 10^{-5} \text{ cm s}^{-1}$, which is within the range commonly

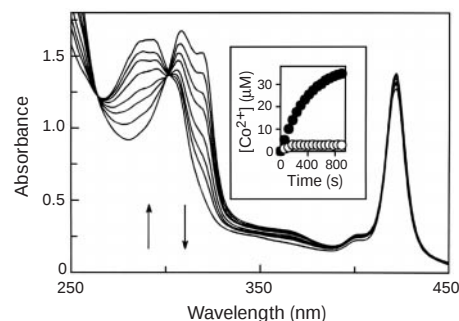


Figure 2 Changes in absorption spectra that accompany illumination of the complete reaction system (Fig. 1). The absorption at 420 nm is the ZnTPPS^{4-} Soret band¹⁰ and the broad bands at 308 nm and at 290 nm are due to $\text{Co}(\text{bpy})_3^{3+}$ and $\text{Co}(\text{bpy})_3^{2+}$, respectively⁵. Conditions are as follows: $[\text{DHP}] = 4$ mM, $[\text{D}(\text{SH})_2] = 500 \mu\text{M}$, $[\text{TPP}^+] = 0.15 \mu\text{M}$ (when present), $[\text{ZnTPPS}^{4-}] = 2.5 \mu\text{M}$, $[\text{Co}(\text{bpy})_3^{3+}] = 47 \mu\text{M}$; excitation at 420 nm using a narrow-bandwidth interference filter⁴. The time course of $\text{Co}(\text{bpy})_3^{2+}$ accumulation, determined from the $\text{Co}(\text{bpy})_3^{3+} - \text{Co}(\text{bpy})_3^{2+}$ difference spectral maximum at 320 nm ($\Delta\epsilon = 2.8 \times 10^4 \text{ M}^{-1} \text{ cm}^{-1}$)⁵, is shown in the inset—filled circles, TPP^+ present; open circles, TPP^+ absent.

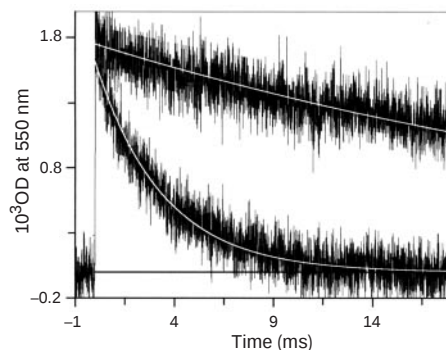


Figure 3 Decay of transient absorption at 550 nm following 355-nm excitation of TPP^+ in vesicular suspensions. Conditions as follows: $[\text{DHP}] = 4$ mM, $[\text{TPP}^+] = 3 \mu\text{M}$; $[\text{Co}(\text{bpy})_3^{3+}] = 0$ (upper trace) or $47 \mu\text{M}$ (lower trace); excitation energy, 150 mJ per pulse. OD, optical density.

measured for permeabilities of small polar non-electrolytes across bilayer membranes¹⁶.

The overall behaviour of this system is consistent with the general reaction features outlined in Fig. 1, wherein net photoinduced one-electron reduction of TPP⁺ is followed by transmembrane diffusion of TPP⁰ to the vesicle interior where it reduces Co(bpy)₃³⁺, reforming TPP⁺; subsequent hydrolysis of TPP⁺ with release of a proton gives TPPD. Diffusion of TPPD to the external solution is then followed by dehydration to TPP⁺ with consumption of a proton to close the cycle of TPP⁺ transformations. Under the conditions of Fig. 2, the amount of Co(bpy)₃³⁺ reduced exceeded the TPP⁺ concentration by more than 23 times, indicating that in this experiment, on average, each TPP⁺ underwent 23 cycles of electron/hydroxide ion counter-transport from electron donors outside the vesicles to acceptors within the vesicles. In general, quantum yields for Co(bpy)₃³⁺ reduction increased with increasing TPP⁺ concentrations between 2 and 20 μM and were independent of the concentration of dithiothreitol (HSCH₂(CHOH)₂CH₂SH, D(SH)₂) when it was above 200 μM; for the conditions given in Fig. 2, ψ = 0.25. Under these conditions, only ~30% of the ³ZnTPPS⁴⁻ ions undergo oxidative quenching by TPP⁺. Given that the intersystem crossing quantum yield for singlet-excited ZnTPPS⁴⁻ is 0.86 (ref. 11), it follows that the transmembrane reduction step occurs with unitary quantum efficiency. Assuming E⁰(Co(bpy)₃^{3+/2+}) = 0.34 V (ref. 17), E⁰(DS₂⁻, 2H⁺/D(SH)₂) ≈ 1.1 V (at pH 5, estimated from data in ref. 18), and the transmembrane potential contributed by the imposed proton gradient is 0.18 V, the electroneutral one-electron photocatalysed reaction, that is, D(SH)₂(o) + Co(bpy)₃³⁺(i) → DS₂H(o) + Co(bpy)₃²⁺(i) + H⁺(i) (Fig. 1) is endothermic by 0.6 V. (The symbols i and o refer to internally and externally localized species, respectively.) Because the DS₂H radical is highly unstable with respect to disproportionation¹⁸, the overall reaction: $\frac{1}{2}$ D(SH)₂(o) + Co(bpy)₃³⁺(i) → $\frac{1}{2}$ DS₂(o) + Co(bpy)₃²⁺(i) + H⁺(i) is exothermic by ~1 V. Based on a porphyrin triplet-state energy of 1.61 V and E⁰(ZnTPPS^{3-/4-}) = 0.87 V (ref. 11), the energetic efficiency of transmembrane charge separation, that is, ³ZnTPPS⁴⁻(o) + Co(bpy)₃³⁺(i) + H⁺(o) → ZnTPPS³⁻(o) + Co(bpy)₃²⁺(i) + H⁺(i), is ~22%. Consistent with the proposed mechanism, practically no reaction was observed when either the internal volume was adjusted to pH 5.0 or the external volume was adjusted to pH 8.0. Conditions in the former case preclude hydrolysis to the membrane-permeable TPPD form and, in the latter case, dehydration of TPPD to the redox-active TPP⁺ form. Net cyclic photoreduction of Co(bpy)₃³⁺ did occur when both inner and outer aqueous phases were at pH 6.5, albeit at a lower overall efficiency. Thus, a transmembrane pH gradient was not required for reactivity.

Achievement of cyclic charge transport allows development of vesicle-based photocatalytic systems for generation of useful chemicals. The permeability of H₂O across bilayer membranes is high¹⁶, consequently, the net effect of e⁻/OH⁻ transport is delivery of both an electron and a proton to the aqueous core. Such an arrangement is optimally suited to continuous photoproduction of H₂; for example, by replacing the sacrificial electron acceptor (Co(bpy)₃³⁺) with a catalyst for H⁺ reduction. □

Methods

DHP small unilamellar vesicles containing occluded Co(bpy)₃³⁺ were prepared in 20 mM Tris/Cl, pH 8.0, by sonication and column chromatography¹⁷, then mixed when needed with an equal volume of 100 mM sodium acetate, pH 5.0; TPP⁺, ZnTPPS⁴⁻ and D(SH)₂ were subsequently added as required from concentrated stock solutions. Before exposure to light, the solutions were made anaerobic by bubbling with Ar. Absorption spectra and formation and decay kinetics of reaction intermediates were measured at ambient temperature with a Nd:YAG transient spectrophotometer by using the second (532 nm) or third (355 nm) harmonic as excitation source⁴.

Received 25 March; accepted 22 October 1999.

1. Hurst, J. K. in *Kinetics and Catalysis in Microheterogeneous Systems* (eds Grätzel, M. & Kalyanasundaram, K.) 183–226 (Surfactant Science Series, Vol. 38, Marcel Dekker, New York, 1991).

2. Lyman, S. V., Parmon, V. N. & Zamarayev, K. I. in *Photoinduced Electron Transfer III* (ed. Mattay, J.) 1–66 (Topics in Current Chemistry, Vol. 159, Springer, Berlin, 1991).
3. Robinson, J. N. & Cole-Hamilton, D. J. Electron transfer across vesicle bilayers. *Chem. Soc. Rev.* **20**, 49–94 (1991).
4. Lyman, S. V., Khairutdinov, R. F., Soldatenkova, V. A. & Hurst, J. K. N-Alkylpyridinium-mediated photoinduced charge separation across dihexadecyl phosphate vesicle membranes. *J. Phys. Chem. B* **102**, 2811–2819 (1998).
5. Lyman, S. V. & Hurst, J. K. Electrostatic and electroneutral pathways for methyl viologen-mediated transmembrane oxidation-reduction across dihexadecylphosphate vesicular membranes. *J. Phys. Chem.* **98**, 989–996 (1994).
6. Hammarstrom, L., Almgren, M. & Norrby, T. Transmembrane electron transfer mediated by a viologen: a mechanism involving diffusion of doubly reduced viologen formed by disproportionation of viologen radical. *J. Phys. Chem.* **96**, 5017–5024 (1992).
7. Futami, A., Hurt, E. & Hauska, G. Vectorial redox reactions of physiological quinones. 1. Requirements of a minimal length of the isoprenoid side chain. *Biochim. Biophys. Acta* **547**, 583–596 (1979).
8. Berson, J. A. Ring-chain tautomerism of pyrylium pseudo-bases. *J. Am. Chem. Soc.* **74**, 358–360 (1952).
9. Williams, A. The hydrolysis of pyrylium salts. Kinetic evidence for hemiacetal intermediates. *J. Am. Chem. Soc.* **93**, 2733–2737 (1971).
10. Tricot, Y.-M. et al. Dihexadecylphosphate vesicle dispersions—preparation, physical properties, and interactions with cationic components used in the solar photolysis of water. *J. Colloid Interface Sci.* **97**, 380–391 (1984).
11. Kalyanasundaram, K. & Neumann-Spallart, M. Photophysical and redox properties of water-soluble porphyrins in aqueous media. *J. Phys. Chem.* **86**, 5163–5169 (1982).
12. Wintgens, V., Poulquien, J., Kossanyi, J. & Heitz, M. Reduction of pyrylium salts: study by ESR and visible spectroscopy of the reversible dimerization of pyranil radical. *Nouv. J. Chem.* **10**, 345–350 (1986).
13. Neta, P. One-electron transfer reactions involving zinc and cobalt porphyrins in aqueous solutions. *J. Phys. Chem.* **85**, 3678–3684 (1981).
14. Kuriyama, Y., Arai, T., Sakuragi, H. & Tokumaru, K. Effects of sensitizer spin multiplicity on electron transfer-initiated isomerization of *cis*-stilbene. *Chem. Lett.* 1193–1196 (1988).
15. Dickinson, T. & Wynne-Jones, W. F. K. Mechanism of Kolbe's electrosynthesis. *Trans. Faraday Soc.* **58**, 382–387 (1962).
16. Walter, A. & Gutknecht, J. Permeability of small nonelectrolytes through lipid bilayer membranes. *J. Membr. Biol.* **90**, 207–217 (1986).
17. Farina, R. & Wilkins, R. G. Electron transfer rate studies of a number of cobalt(II)-cobalt(III) systems. *Inorg. Chem.* **7**, 514–518 (1968).
18. Surdhar, P. S. & Armstrong, D. A. Reduction potentials and exchange reactions of thiyl radicals and disulfide anion radicals. *J. Phys. Chem.* **91**, 6532–6537 (1987).

Acknowledgements

This work was supported by the Office of Basic Energy Sciences, US Department of Energy.

Correspondence and requests for materials should be addressed to J.K.H. (e-mail: hurst@wsu.edu).

Warming of the tropical Atlantic Ocean and slowdown of thermohaline circulation during the last deglaciation

Carsten Rühlemann*, Stefan Mulitza*, Peter J. Müller*, Gerold Wefer* & Rainer Zahn†

* University of Bremen, Department of Geosciences, Klagenfurter Strasse, D-28359 Bremen, Germany

† GEOMAR Research Center for Marine Geosciences, Wischhofstrasse 1-3, D-24148 Kiel, Germany

Evidence for abrupt climate changes on millennial and shorter timescales is widespread in marine and terrestrial climate records^{1–4}. Rapid reorganization of ocean circulation is considered to exert some control over these changes⁵, as are shifts in the concentrations of atmospheric greenhouse gases⁶. The response of the climate system to these two influences is fundamentally different: slowing of thermohaline overturn in the North Atlantic Ocean is expected to decrease northward heat transport by the ocean and to induce warming of the tropical Atlantic^{7,8}, whereas

atmospheric greenhouse forcing should cause roughly synchronous global temperature changes⁹. So these two mechanisms of climate change should be distinguishable by the timing of surface-water temperature variations relative to changes in deep-water circulation. Here we present a high-temporal-resolution record of sea surface temperatures from the western tropical North Atlantic Ocean which spans the past 29,000 years, derived from measurements of temperature-sensitive alkenone unsaturation in sedimentary organic matter. We find significant warming is documented for Heinrich event H1 (16,900–15,400 calendar years BP) and the Younger Dryas event (12,900–11,600 cal. yr BP), which were periods of intense cooling in the northern North Atlantic. Temperature changes in the tropical and high-latitude North Atlantic are out of phase, suggesting that the thermohaline circulation was the important trigger for these rapid climate changes.

In the high northern latitudes, the last deglaciation was punctuated by several brief returns to near-glacial climate^{1,2}. The most prominent of these events was the Younger Dryas cold period. Temperatures over Greenland, as seen in the GRIP ice-core $\delta^{18}\text{O}$ record (Fig. 1a), are essentially controlled by heat released from the surface of the North Atlantic Ocean². This heat is delivered from the warm-water reservoir in the western tropical Atlantic through northward transport of warm upper-layer waters, which are imported to compensate for the southward flow of North Atlantic Deep Water (NADW) formed in the Norwegian–Greenland and Labrador seas. As a result of this cross-equatorial exchange of water, the Southern Hemisphere loses heat⁷. Modelling experiments predict an abrupt Younger Dryas cooling of the surface northern North Atlantic in response to freshwater-induced slowdown of NADW formation and concomitant reduction in northward heat transport. This would be accompanied by a warming in the western tropical North Atlantic and throughout most of the Southern Hemisphere⁸. An asynchronous coupling of the two hemispheres caused by changes of thermohaline circulation is also inferred from temperature records from Arctic and Antarctic ice cores¹⁰ (Fig. 1a and b), and from palaeoceanographic records which display warm surface temperatures in the southern Atlantic during periods of reduced NADW production^{3,11}. Reconstructions of high altitude temperatures¹² and glacier advance¹³ in the northern and southern Andean mountains, on the other hand, indicate that these regions cooled in phase with the Younger Dryas cooling of the high northern latitudes. This evidence has been used to suggest that climates of the two hemispheres were mainly connected through alterations of greenhouse-gas concentrations which result in a nearly synchronous response of surface temperatures due to the rapid mixing of the atmosphere⁹. To distinguish between oceanic and atmospheric processes linking the North Atlantic climate to distant regions, it is necessary to investigate the geographical pattern of surface temperature variability during millennial-scale climate perturbations.

Sediment core M35003-4, located southeast of the island of Grenada (12° 05' N, 61° 15' W; 1,299 m water depth), shows high sedimentation rates of between 10 and 20 cm kyr⁻¹, and is therefore well suited to monitor deglacial changes of the tropical Atlantic's heat reservoir. We determined sea surface temperatures (SST) by using the alkenone method and $\delta^{18}\text{O}$ values of the planktic foraminifer *Globigerinoides ruber* (pink) with an average sampling interval of 300 years (see Methods section). The stratigraphy is based on 12 accelerator mass spectrometry (AMS) ¹⁴C datings (Table 1). For SST reconstructions, we used the simplified unsaturation index of C₃₇ alkenones ($U_{37}^{K'}$; ref. 14). To convert from $U_{37}^{K'}$ values to SST we applied the calibration of Müller *et al.*¹⁵ (T (°C) = ($U_{37}^{K'} - 0.044$) / 0.033) which is based on an extensive suite of core-top samples from the world ocean between 60° N and 60° S. The 1 σ standard deviation (s.d.) of this global calibration is ± 1.5 °C, but reduces to ± 0.7 °C (ref. 16) and even ± 0.2 °C (ref. 17) if more regional data sets from the low-latitude ocean with SSTs ranging between 24 and 29 °C are

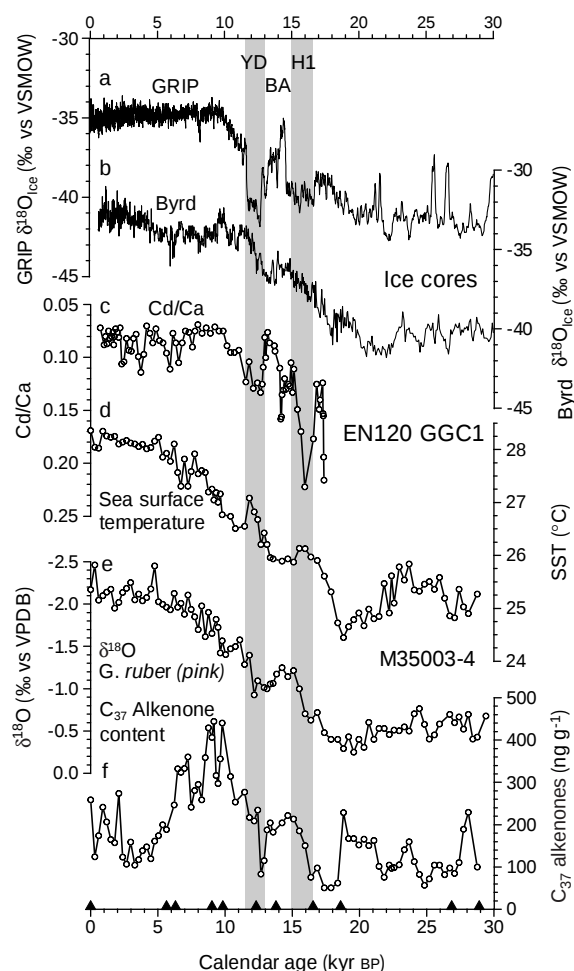


Figure 1 Temperature evolution of the western tropical Atlantic Ocean and of the Arctic and Antarctic during the past 30,000 yr. Shown is a comparison of alkenone-based sea surface temperatures (SST; **d**) and oxygen isotope ratios of the planktic foraminifer *G. ruber* (pink) (**e**) from sediment core M35003-4 with oxygen isotope records from Greenland (GRIP) (**a**) and Antarctica (Byrd) (**b**), and with cadmium/calcium ratios of benthic foraminiferal tests in sediment core EN120 GGC1 (33°40'N/57°37'W; 4,450 m) from the Bermuda rise²⁹ (**c**; axis inverted). **f**, C₃₇ alkenone content of core M35003-4. ¹⁴C AMS control points for M35003-4 are denoted by triangles on lower axis. Age control for the GRIP and Byrd ice cores is after Blunier *et al.*¹⁰. H1, BA and YD denote Heinrich event H1 (16,900 – 15,400 cal. yr BP), the Bölling–Allerød period (15,400 – 12,900 cal. yr BP), and the Younger Dryas period (12,900 – 11,600 cal. yr BP), respectively.

used. A calibration obtained on *Emiliania huxleyi* grown in laboratory cultures¹⁴, where ambient temperatures are exactly known and not derived from generalized ocean atlas values, yields a s.d. of ± 0.6 °C. We thus estimate the absolute error of our temperature reconstructions from a single location in the warm ocean to be less than ± 1 °C. The study of Müller *et al.*¹⁵ shows that $U_{37}^{K'}$ reflects annual mean mixed-layer SST in the low-latitude ocean, and furthermore suggests that the sedimentary record of $U_{37}^{K'}$ is not significantly affected by changes of species assemblage, growth rate of algae or nutrient availability. A comparison of oxic and anoxic samples from turbidite sequences has indicated that even strong diagenesis of alkenones usually does not bias the $U_{37}^{K'}$ by more than 0.03 units (or ± 1.0 °C; ref. 18). We suggest, however, that the redox conditions at the site of core M35003-4 remained fairly stable and that the diagenetic overprint is small, given the low organic-carbon content ranging between 0.3 and 0.7 wt%.

Quantification of glacial SSTs in tropical regions is still controversial, with much debate whether temperatures were similar to

Table 1 ^{14}C dates used to constrain the age–depth model for core M35003-4

Sample	Core depth (cm)	Species	^{14}C AMS age (kyr BP)	Error (yr)	Calendar age (cal. kyr BP)
KIA4693	3	Mixed planktics	0.38	± 30	0
KIA5085	98	<i>G. ruber</i> (pink)	5.30	± 50	5.65
KIA4223	110	<i>G. ruber</i> (white)	5.92	± 40	6.31
KIA5084	163	Mixed planktics	8.48	± 60	9.02
KIA4224	190	<i>G. ruber</i> (white)	9.15	± 50	9.84
KIA4225	225	<i>G. ruber</i> (white)	10.80	± 90	12.30
KIA4696	228	Mixed planktics	10.76	± 80	12.24
KIA4226	257.5	<i>G. ruber</i> (white)	12.22	± 70	13.78
KIA4227	290	<i>G. ruber</i> (white)	14.21	± 90	16.56
KIA4228	310	<i>G. ruber</i> (white)	16.14	± 100	18.60
KIA4229	420	<i>G. ruber</i> (white)	23.23	± 210	26.87
KIA4230	450	<i>G. ruber</i> (white)	25.06	± 260	28.91
KIA4115	463	Mixed planktics	27.03	$-220/+210$	31.06

We have not used the radiocarbon age at 228 cm to constrain the stratigraphy; it simply documents the accuracy of the dating at 225 cm which is particularly critical for our conclusion of a Younger Dryas warming. The two calendar ages are identical within the 1σ error limits (12.43–12.14 cal. kyr for 225 cm, and 12.38–12.08 cal. kyr for 228 cm). The plateau observed in the ^{14}C ages at 225 and 228 cm is probably related to a decrease in the ^{14}C production rate during the course of the Younger Dryas³³.

modern-day values or decreased by 5–6 °C; ref. 19). According to our data, temperatures during the Last Glacial Maximum (21,000 cal. yr BP) were decreased by 3.5 °C (Fig. 1d)—that is, 2 °C below CLIMAP²⁰ estimates derived from foraminiferal assemblages, and 2 °C above SST reconstructions based on $\delta^{18}\text{O}$ and Sr/Ca ratios in Barbados corals located some 250 km northeast of the position of core M35003-4 (ref. 21). Our alkenone-derived temperature difference coincides with recent SST estimates based on a revised foraminiferal temperature equation²², and agrees with the glacial/Holocene SST increase of 2.6 ± 1.3 °C deduced from Mg/Ca ratios in foraminiferal calcite²³. The alkenone estimate is confirmed further by the deglacial $\delta^{18}\text{O}$ decrease measured on tests of the surface-dwelling foraminifer *G. ruber* (pink) from the same sediment core (Fig. 1e). Depending on the applied ice volume correction of 1.0‰ (ref. 24) to 1.2‰ (ref. 25), the glacial/Holocene difference of 1.8‰ equals a temperature shift of 3–4 °C (assuming a decrease in $\delta^{18}\text{O}$ of 0.22‰ per 1 °C temperature increase). The result is consistent with the large ice-age cooling over tropical land areas (5–6 °C) if account is taken of a sea-level lowering of 120 m (ref. 26) and reductions of the tropical rainforest area²⁷.

Our alkenone-based SST estimates suggest that during the early deglacial, between 19,000 and 16,000 cal. yr BP, western tropical Atlantic temperatures increased by ~ 1.5 °C followed by a temperature plateau between 15,000 and 13,000 cal. yr BP and a second warming step (1.2 °C) centred around 12,000 cal. yr BP (Fig. 1d). SSTs increased steadily from 10,000 cal. yr BP (26.5 °C) up to present-day values of 28.3 °C. Although the alkenone-derived SST evolution is not expressed in the planktic $\delta^{18}\text{O}$ record between 14,000 and 10,000 cal. yr BP, it is generally consistent with SST estimates from the Modern Analogue Technique (MAT) using planktic foraminiferal assemblages along the same core²⁸. These data confirm the alkenone-derived SST increase during Heinrich event H1 and during the Younger Dryas, but suggest significant cooling of ~ 2 °C between 15,000 and 13,000 cal. yr BP, coinciding with the Bølling–Allerød period²⁸. The $\delta^{18}\text{O}$ values of *G. ruber* (pink) increase by some 0.3‰ from the Bølling–Allerød to the Younger Dryas, whereas alkenone-derived SSTs are already starting to rise. The slight $\delta^{18}\text{O}$ increase may have been caused by a change in the evaporation–precipitation balance due to a northward shift of the Intertropical Convergence Zone, a reduction in the Orinoco River freshwater discharge, or changes in the outflow pattern. The $\delta^{18}\text{O}$ –salinity relationship that we obtained from 94 surface-water samples of the eastern and southern Caribbean Sea is $\delta^{18}\text{O} = 0.3S - 9.56$, where S is salinity (S.M., unpublished data). That is, an increase in salinity of just 1‰ would result in a $\delta^{18}\text{O}$ decrease of about 0.3‰.

It is significant that the SST variations in the western tropical Atlantic between 17,000 and 10,000 cal. yr BP are in antiphase to

those of the northern North Atlantic as inferred from the Greenland GRIP ice-core $\delta^{18}\text{O}$ record (Fig. 1a), but are similar to the Antarctic Byrd ice-core temperature record (Fig. 1b). During Heinrich event H1 and the Younger Dryas, SSTs increased in the tropical Atlantic whereas temperatures over Greenland sharply dropped by several degrees centigrade. During the Bølling–Allerød period, which was characterized by warm temperatures over Greenland, SSTs in the western tropical Atlantic were cooler than during the Younger Dryas. This thermal antiphasing between the tropical Atlantic and Greenland during the deglacial warming holds important clues to the response of the tropical Atlantic surface waters to high-latitude ocean and climate processes.

A comparison of our temperature reconstruction (Fig. 1d) with a high-resolution record of Cd/Ca ratios in benthic foraminiferal tests from the Bermuda rise²⁹ (Fig. 1c) strongly suggests that changes of the thermohaline circulation have modulated the deglacial temperature evolution in the western tropical Atlantic. The Cd/Ca ratios trace deep-water nutrient variability; this variability is linked to changes in water mass contributions from low-nutrient (low Cd/Ca) North Atlantic sources and high-nutrient (high Cd/Ca) Antarctic sources. The Cd/Ca maxima during Heinrich event H1 and the Younger Dryas indicate significant slowdown of NADW formation during these periods. These slowdowns were caused by the injection of large volumes of fresh water into the northern North Atlantic, resulting in decreased sea surface salinity and diminished deep convective overturning^{6,8}. As a consequence, the compensating import of warm water masses from the tropics was reduced and atmospheric temperatures over Greenland (GRIP) decreased whereas SSTs in the western tropical Atlantic rose. The sparse palaeoceanographic evidence that is currently available suggests that during times of reduced NADW formation, surface water warming was indeed widespread in the South Atlantic^{3,11,30}. In contrast to the western Atlantic, the eastern low-latitude North Atlantic exhibits a surface temperature evolution that is synchronous with that of the northern North Atlantic. High-resolution alkenone records from off northwest Africa (20° N) indicate rapid surface water cooling during Heinrich event H1 and during the Younger Dryas³¹. Southward advection of cold surface water with the Canary Current has transferred the high-latitude climate signal during these ice-rafting events to the eastern subtropical North Atlantic³¹. Hence a geographical pattern emerges in which the western part of the tropical North Atlantic and the South Atlantic warmed up in response to slowdown of thermohaline overturn, while the eastern and northern North Atlantic cooled.

Our results agree with evidence for the asynchronous coupling of Arctic and Antarctic temperatures as deduced from the Greenland and Antarctic ice-core records¹⁰. The antiphasing of deglacial, millennial-scale temperature fluctuations suggests that sudden changes in the thermohaline circulation have profound consequences for the temperature distribution, at least in the northern and tropical Atlantic Ocean. It is likely that the increased temperature difference between high and low northern latitudes increased the atmospheric pressure gradient, and thus caused enhanced intensities of the northeast trade wind. This hypothesis, however, needs to be tested by coupled ocean–atmosphere models.

Methods

Alkenones

Samples of 2–4 g (depending on the organic carbon content) freeze-dried and homogenized sediment were mixed with an internal standard (squalane; $\text{C}_{30}\text{H}_{62}$), and extracted with a mixture of successively less polar, twice-distilled methanol and methylene chloride (CH_3OH , $\text{CH}_3\text{OH}/\text{CH}_2\text{Cl}_2$ 1:1, CH_2Cl_2) by ultrasonication using a UP 200H sonic disruptor probe (Hielscher GmbH, 200 W, amplitude 105 μm , pulse 0.5 s). After centrifuging the extracts were combined, desalted with de-ionized water, dried with preheated Na_2SO_4 , and rotary evaporated to dryness. The residues were redissolved in CH_2Cl_2 and cleaned by elution through a commercial silica cartridge (silica Bond Elute SPE column, Varian). To eliminate interference with wax esters, the clean extracts were hydrolysed with 0.1 N KOH in methanol (90/10 $\text{CH}_3\text{OH}/\text{H}_2\text{O}$) at 80 °C for 2 h, and the neutral fraction

containing the alkenones was obtained by partitioning into hexane. Finally, the extracts were concentrated under N_2 and taken up in 50 μ l CH_2Cl_2 . Gas chromatography was performed using a HP 5890 series II gas chromatograph equipped with a cold on-column injector, a 60 m $\times$ 0.25 mm $\times$ 0.1 μ m nonpolar fused silica capillary column DB-5MS (J&W) and a flame ionization detector (with helium as carrier gas). The oven temperature was programmed to give 40–200 °C at 15 °C min⁻¹, 200–250 °C at 5 °C min⁻¹, and 250–300 °C at 3 °C min⁻¹ and the final temperature maintained for 35 min. Analytical precision ($\pm 1\sigma$), based on duplicates and multiple extractions of a sediment used as a laboratory-internal reference sample, was better than 0.01 U_{37}^K units (or 0.3 °C). To avoid analytical bias due to selective adsorption of C_{37} alkenones at low concentrations³² (Fig. 1e) we injected a close range of total alkenone amounts (in most cases 5–15 ng) into the gas chromatograph.

Stable oxygen isotopes

The faunal samples were washed over 150- and 63- μ m sieves and dried in an oven at 60 °C. From each sample, 15–30 specimens of *G. ruber* (pink) (350–400 μ m) were picked out for isotope analyses. The isotopic composition of the foraminiferal shells was measured using a Finnigan MAT 251 mass spectrometer equipped with an automatic carbonate preparation device. Internal precision, based on replicates of a limestone standard, was better than 0.07‰.

Radiocarbon stratigraphy

Nine ages were determined on monospecific samples of *G. ruber* (white) and four ages on a mixture of *G. ruber* (white and pink), *Globigerinoides sacculifer* and *Orbulina universa*. The ^{14}C ages were corrected for a reservoir age of 400 years (ref. 33); they were then converted to calendar ages using the program CALIB 3.0.3c (ref. 34) for ages <20,000 yr and using the equation given in the caption of Figure 3 of ref. 4 for ages >20,000 yr. For the deglacial, we conservatively estimate the calendar year chronology to be good within ± 500 yr. Before about 25 cal. kyr BP there are only a limited number of calibration points, thus the assumed correction has an uncertainty within at least $\pm 1,500$ yr.

Received 20 April; accepted 12 October 1999.

- Dansgaard, W. *et al.* Evidence for general instability of past climate from a 250-kyr ice-core record. *Nature* **364**, 218–220 (1993).
- Bond, G. *et al.* Correlations between climate records from North Atlantic sediments and Greenland ice. *Nature* **365**, 143–147 (1993).
- Charles, C. D., Lynch-Stieglitz, J., Ninnemann, U. S. & Fairbanks, R. G. Climate connections between the hemisphere revealed by deep sea sediment core/ice core correlations. *Earth Planet. Sci. Lett.* **142**, 19–27 (1996).
- Bard, E., Rostek, F. & Sonzogni, C. Interhemispheric synchrony of the last deglaciation inferred from alkenone palaeothermometry. *Nature* **385**, 707–710 (1997).
- Broecker, W. S., Peteet, D. M. & Rind, D. Does the ocean–atmosphere system have more than one stable mode of operation. *Nature* **315**, 21–26 (1985).
- Broecker, W. S. Massive iceberg discharges as triggers for global climate change. *Nature* **372**, 421–424 (1994).
- Crowley, T. J. North Atlantic Deep Water cools the southern hemisphere. *Paleoceanography* **7**, 489–497 (1992).
- Manabe, S. & Stouffer, R. J. Coupled ocean–atmosphere model response to freshwater input: Comparison to Younger Dryas event. *Paleoceanography* **12**, 321–336 (1997).
- Manabe, S., Stouffer, R. J., Spelman, M. J. & Bryan, K. Transient response of a coupled ocean–atmosphere model to gradual changes of atmospheric CO_2 . Part I: Annual mean response. *J. Clim.* **4**, 785–818 (1991).
- Blunier, T. *et al.* Asynchrony of Antarctic and Greenland climate change during the last glacial period. *Nature* **394**, 739–743 (1998).
- Vidal, L. *et al.* Link between the North and South Atlantic during the Heinrich events of the last glacial period. *Clim. Dyn.* (in the press).
- Thompson, L. G. *et al.* A 25,000-year tropical climate history from Bolivian ice cores. *Science* **282**, 1858–1864 (1998).
- Lowell, T. V. *et al.* Interhemispheric correlation of Late Pleistocene glacial events. *Science* **269**, 1541–1549 (1995).
- Prahl, F. G. & Wakeham, S. G. Calibration of unsaturation patterns in long-chain ketone compositions for paleotemperature assessment. *Nature* **330**, 367–369 (1987).
- Müller, P. J., Kirst, G., Ruhland, G., von Storch, I. & Rosell-Melé, A. Calibration of the alkenone paleotemperature index U_{37}^K based on core-tops from the eastern South Atlantic and the global ocean (60°N–60°S). *Geochim. Cosmochim. Acta* **62**, 1757–1772 (1998).
- Sonzogni, C. *et al.* Temperature and salinity effects on alkenone ratios measured in surface sediments from the Indian Ocean. *Quat. Res.* **47**, 344–355 (1997).
- Pelejero, C. & Grimalt, J. O. The correlation between U_{37}^K index and sea surface temperatures in the warm boundary: The South China Sea. *Geochim. Cosmochim. Acta* **61**, 4789–4797 (1997).
- Prahl, F. G., de Lange, G. J., Lyle, M. & Sparrow, M. A. Post-depositional stability of long-chain alkenones under contrasting redox conditions. *Nature* **341**, 434–437 (1989).
- Broecker, W. Glacial climate in the tropics. *Science* **272**, 1902–1904 (1996).
- CLIMAP Project Members *Seasonal Reconstruction of the Earth's Surface at the Last Glacial Maximum* (Map & Chart Ser. MC-3, Geol. Soc. Am., Boulder, Colorado, 1981).
- Guilderson, T. P., Fairbanks, R. G. & Rubenstone, J. L. Tropical temperature variations since 20,000 years ago: Modulating interhemispheric climate change. *Science* **263**, 663–665 (1994).
- Mix, A. C., Morey, A. E., Pisias, N. G. & Hostetler, S. W. Foraminiferal fauna estimates of paleotemperature: Circumventing the no-analog problem yields cool ice age tropics. *Paleoceanography* **14**, 350–359 (1999).
- Hastings, D. W., Russell, A. D. & Emerson, S. R. Foraminiferal magnesium in *Globigerinoides sacculifer* as a paleotemperature proxy. *Paleoceanography* **13**, 161–169 (1998).
- Schrag, D. P., Hampt, G. & Murray, D. W. Pore fluid constraints on the temperature and oxygen

- isotopic composition of the glacial ocean. *Science* **272**, 1930–1932 (1996).
- Fairbanks, R. G. A 17,000-year glacio-eustatic sea level record: influence of glacial melting rates on the Younger Dryas event and deep-ocean circulation. *Nature* **342**, 637–642 (1989).
- Betts, A. K. & Ridgway, W. Tropical boundary layer equilibrium in the last ice age. *J. Geophys. Res.* **97**, 2529–2534 (1992).
- Crowley, T. J. & Baum, S. K. Effect of vegetation on an ice-age climate model simulation. *J. Geophys. Res.* **102**, 16463–16480 (1997).
- Hüls, M. & Zahn, R. Millennial-scale SST variability in the western subtropical North Atlantic from planktonic foraminiferal census counts. *Paleoceanography* (submitted).
- Boyle, E. A. & Keigwin, L. North Atlantic thermohaline circulation during the past 20,000 years linked to high-latitude surface temperature. *Nature* **330**, 35–40 (1987).
- Arz, H. W., Pätzold, J. & Wefer, G. The deglacial history of the western tropical Atlantic as inferred from high resolution stable isotope records off northeastern Brazil. *Earth Planet. Sci. Lett.* **167**, 105–117 (1999).
- Zhao, M., Beveridge, N. A. S., Shackleton, N. J., Sarnthein, M. & Eglinton, G. Molecular stratigraphy of cores off northwest Africa: Sea surface temperature history over the last 80 ka. *Paleoceanography* **10**, 661–675 (1995).
- Villanueva, J. & Grimalt, J. O. Gas chromatographic tuning of the U_{37}^K paleothermometer. *Anal. Chem.* **69**, 3329–3332 (1997).
- Hughen, K. A. *et al.* Deglacial changes in ocean circulation from an extended radiocarbon calibration. *Nature* **391**, 65–68 (1998).
- Stuiver, M. & Reimer, P. J. Extended ^{14}C data base and revised CALIB 3.0 ^{14}C age calibration program. *Radiocarbon* **35**, 215–230 (1993).
- Johnsen, S. J., Dansgaard, W., Clausen, H. B. & Langway, C. C. Oxygen isotope profiles through the Antarctic and Greenland ice sheets. *Nature* **235**, 429–434 (1972).

Acknowledgements

We thank A. Rosell-Melé, R. Schneider and A. Paul for comments on the manuscript, M. Hüls for discussions, and P. Grootes and staff of the Leibnitz-Labor in Kiel for providing AMS ^{14}C datings. This work was supported by the Deutsche Forschungsgemeinschaft.

Correspondence and requests for materials should be addressed to C.R. (e-mail: ruehl@uni-bremen.de). All data are also available from the PANGAEA database (<http://www.pangaea.de/Projects/GeoB/CRuehlemann>).

The remelting of hydrothermally altered peridotite at mid-ocean ridges by intruding mantle diapirs

Mathieu Benoit*, Georges Ceuleneer & Mireille Polvé

Centre National de la Recherche Scientifique, UMR 5562 & 5563, Observatoire Midi-Pyrénées, 14 avenue Edouard Belin, 31400 Toulouse, France

Most gabbroic cumulates found at ocean spreading centres are thought to have been generated by the fractional crystallization of melts with the composition of mid-ocean ridge basalt (MORB)¹. There are exceptions, however, including some cumulates which appear to have come from melts that contain more silica than MORB and are much more depleted in the incompatible elements (those elements that do not readily substitute into the main mineral phases)². These unusual rocks bear witness to relatively deep petrological processes that are not accessible through the study of melts erupted on the sea floor, and their origin is still debated. Fortunately, the same lithologies can be studied in detail in ophiolites (sections of oceanic crust accreted to a continent). In a fossil mantle diapir of the Oman ophiolite^{3–4}, we have observed the same dichotomy between a suite of 'normal', MORB-type, cumulates ('N-cumulates') and a suite of cumulates issued from silica-enriched but incompatible-element-depleted melts ('D-cumulates'). While the N-cumulates crystallized inside the diapir, the D-cumulates occur essentially as intrusions surrounding the diapir. The combination of silica enrichment, extreme depletion in incompatible elements, and seawater isotopic signa-

* Present address: CNRS, UMR 6538, IUEM, Place N. Copernic, 29280 Plouzané, France.

Table 1 Mineralogical and geochemical characteristics of D-gabbronorites, D-pyroxenites and N-gabbronorites.

TYPE	D-pyroxenites				D-gabbronorites			
Trace element measurements	CPX SIMS	WR ICP-MS	WR ICP-MS	WR ICP-MS	CPX SIMS	CPX SIMS	CPX SIMS	CPX SIMS
Sample	92 OG 55	92 OG 250C	92 OG 246	92 OG 281	95 M 150A	95 M 70	95 M 115A	95 M 123A
⁸⁷ Sr/ ⁸⁶ Sr (100 Myr)	0.706166±17	0.703573±12	0.706461±13	0.705909±13	0.706593±12	0.708324±16	0.706405±12	0.704951±10
Leachate		0.705572±12				0.707238±27	0.703793±23	
εNd (100 Myr)	9.85	8.25	7.51	6.56	4.93	9.03	7.62	6.39
Anorthite content	96	96	96	92	94	95	91	94
Mg number cpx (default opx)	89	87	86	82	82	86	78	83
TiO ₂ cpx (oxide %)	0.06	0.11	0.17	0.02	0.21	0.13	0.48	0.19
Na ₂ O cpx (oxide %)	0.12	0.15	0.16	0.04	0.16	0.05	0.22	0.16
[Nd] (p.p.m.)	0.1581	0.5262	0.2373	0.4878	0.2051	0.1416	1.7902	0.4636
[Yb] (p.p.m.)	0.3284	0.5631	0.3826	1.0085	0.4845	0.8068	1.7413	0.6051
[Sr] (p.p.m.)	3.77					0.589	6.49	5.41
Plagioclase	<1	7	29	31	37	39	30	28
Clinopyroxene	59	59	34	39	20	23	30	30
Orthopyroxene	41	34	32	26	41	38	35	42
Amphibole	0	0	0	9	2	0	5	0
Olivine	0	0	5	0	0	0	0	0
[Nd] N eq. Liq.	1.53	5.76	4.48	10.40	4.26	1.33	17.63	4.96
[Yb] N eq. Liq.	3.47	7.11	7.88	19.17	10.86	8.01	18.99	6.40
[Nd]/[Yb] N eq. Liq.	0.39	0.81	0.57	0.54	0.39	0.13	0.90	0.61

TYPE	N-gabbronorites							
Trace element measurements	WR ICP-MS	WR ICP-MS	WR ICP-MS	CPX SIMS	CPX SIMS	WR ICP-MS	WR ICP-MS	WR ICP-MS
Sample	92 OG 278	94 M 615	95 M 34A	95 M 116A	95 M 106	95 M 140A	95 M 169A	89 OG 23B
⁸⁷ Sr/ ⁸⁶ Sr (100 Myr)	0.702824±10	0.703375±12	0.703599±23	0.703103±9	0.704464±17	0.702818±20	0.703353±13	0.703373±13
Leachate			0.703872±17					
εNd (100 Myr)	8.48	8.81	8.50	7.82	7.52	8.65	8.55	9.41
Anorthite content	85	78	79	88	86	75	80	79
Mg number cpx (default opx)	87	78	81	87	90	85	85	81
TiO ₂ cpx (oxide %)	0.68	0.59	0.62	0.59	0.49	0.62	0.69	0.55
Na ₂ O cpx (oxide %)	0.35	0.32	0.36	0.26	0.26	0.39	0.37	0.29
[Nd] (p.p.m.)	0.8301	2.0275	1.1061	2.9174	1.3106	0.6876	1.3147	0.9594
[Yb] (p.p.m.)	0.4261	1.2848	0.6146	1.5067	0.6572	0.4396	0.6071	0.8483
[Sr] (p.p.m.)				12.87	13.49			
Plagioclase	35	46	54	42	48	45	48	42
Clinopyroxene	25	32	24	24	29	29	10	44
Orthopyroxene	11	10	2	7	2	6	10	8
Amphibole	2	4	0	0	0	3	12	0
Olivine	27	8	18	27	21	17	20	6
[Nd] N eq. Liq.	15.85	28.90	27.22	28.19	11.11	10.68	39.62	15.23
[Yb] N eq. Liq.	12.06	28.08	24.52	15.93	6.95	11.42	24.12	18.02
[Nd]/[Yb] N eq. Liq.	1.31	0.96	1.11	1.55	1.60	0.94	1.64	0.84

⁸⁷Sr/⁸⁶Sr isotopic composition was measured for leached clinopyroxene (see Fig. 3 legend) except for samples 94 M 615 and 95 M 169A on whole rock powder. TiO₂ concentration, and Mg number of clinopyroxene, together with anorthite content (An%) of plagioclase, were determined using electron microprobe analysis. The Nd and Yb concentrations were determined either with ion-probe on clinopyroxene or on whole rock (ICP-MS). Sr concentrations for clinopyroxene are given only for samples measured with ion-probe. Modal compositions were calculated from whole rock and phase compositions and checked by point-counting (reported here). Trace elements for calculated equilibrium liquids were determined using partition coefficients^{20–24} and modal compositions for each gabbronorite. Liquid concentrations are calculated in equilibrium with clinopyroxene. Partition coefficient for Nd: $K_{d_{\text{cpx}}} = 0.1873$, $K_{d_{\text{plg}}} = 0.03$, $K_{d_{\text{opx}}} = 0.037$, $K_{d_{\text{sl}}} = 0.0002$. Partition coefficient for Yb: $K_{d_{\text{cpx}}} = 0.43$, $K_{d_{\text{plg}}} = 0.0085$, $K_{d_{\text{opx}}} = 0.2605$, $K_{d_{\text{sl}}} = 0.00522$. N: chondrite normalized.

ture indicates that the D-cumulates were formed by the remelting at low pressure of hydrated residual peridotites left after MORB extraction at the ridge axis. The distribution of the D-cumulates relative to the N-cumulates suggests that such depleted melts are produced episodically at ridge axes when the lithospheric mantle is reheated by a new diapiric pulse.

Most MORBs have suffered variable degrees of fractional crystallization before their eruption on the sea floor. Crystallization behaviour of MORBs (olivine tholeiites) at low pressure (< 0.3 MPa) is well established empirically and experimentally^{1,5–7}. It leads to a characteristic sequence of cumulus lithologies, including dunites-troctolites, olivine gabbros, oxide-gabbronorites and plagiogranites as the temperature decreases. We have shown recently that the mafic cumulates that crystallized in porous flow channels and dykes in a fossil mantle diapir (Maqсад area of Oman) present a clear distribution of their mineralogy that is globally consistent with this crystallization sequence and suggests decreasing temperature away from the centre of the diapir^{3,4}. In detail, however, some of these cumulates do not fit into this fractional crystallization scheme. This is the case for some of the gabbronorites, websterites and plagiogranites that massively intrude into the peridotites at the periphery of the diapir and are exceptionally found inside the diapir, associated with fault zones^{3,8}. It is clear that a single process cannot account for the petrogenesis of all Maqсад cumulates. This led us to reconsider in greater detail the petrogenesis of Maqсад gabbronor-

ites and associated lithologies, using new major element (microprobe), trace element (ICP-MS, ion probe) and isotopic data.

Petrographically and geochemically the Maqсад gabbronorites belong to two clearly distinct families (Table 1), one with 'normal' mineralogical and chemical compositions for MORB cumulates ('N-gabbronorites') and the other with much more depleted chemical compositions and 'abnormal' mineralogy for MORB cumulates ('D-gabbronorites').

The N-gabbronorites occur in swarms of centimetre- to decimetre-thick sub-vertical dykes lying approximately parallel to the azimuth of the sheeted dyke complex (inferred to be the azimuth of the ridge axis). They have fine-grained textures (grain size of ~1 mm) with local development of coarse-grained margins at the dyke walls. They are composed of cumulus plagioclase (~45%), clinopyroxene (~27%) and olivine (~18%). Orthopyroxene occurs in minor amounts (≤11%) and has post-cumulus characteristics (anhedral and interstitial); this is associated with brown hornblende developing at its expense and with abundant Fe–Ti oxides.

The D-gabbronorites are mainly found in giant dykes, up to several metres in thickness, with random orientation. Textural variations in these dykes, from fine-grained to pegmatitic, define a crude layering parallel to their margins. They are generally associated with intrusive bodies of pegmatitic websterites reaching several hundred metres in size. Orthopyroxene is a major phase (~37%); olivine content is low (0–5%). Orthopyroxene forms

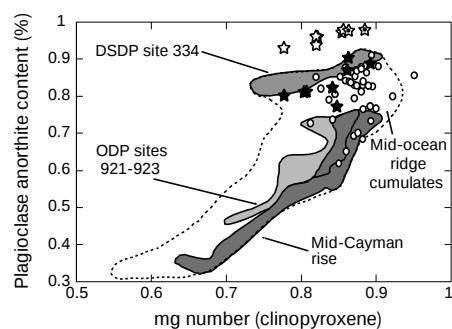


Figure 1 Correlation between Mg number (clinopyroxene) and An% (plagioclase) for the mafic cumulates trapped within the Maqsd harzburgites³. Circles, olivine gabbros and clinopyroxene-troctolites; filled stars, N-gabbro-norites; open stars, D-gabbro-norites; open, dotted stars, D-websterites. DSDP site 334 field from ref. 2, ODP site 921–923 field from ref. 6 and mid-Cayman rise field from ref. 7.

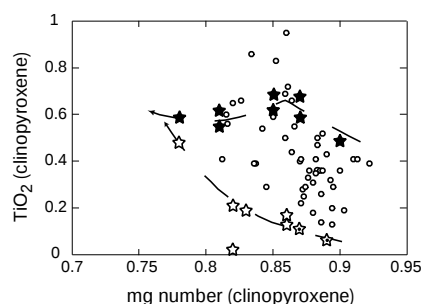


Figure 2 TiO₂ content of clinopyroxene versus Mg number. Symbols as in Fig. 1. Arrows show the contrasting differentiation trends of D- and N-gabbro-norites. The decrease in TiO₂ concentrations of the most evolved N-gabbro-norites is related to hornblende and Fe–Ti oxide crystallization. Na₂O contents (not shown) present the same trends.

euhedral to subhedral phases. Plagioclase and clinopyroxene are mostly interstitial. Brown hornblende is generally absent and the oxide content is very low (<1%).

The contrast in phase chemistry between D- and N-gabbro-norites is illustrated by two examples. While plagioclase in N-gabbro-norites has anorthite content (An%) that falls in the range of mid-ocean ridge cumulates content, plagioclase in all D-gabbro-norites have highly calcic plagioclase (An% 90–96) (Fig. 1). The incompatible element content of pyroxenes (like TiO₂ in clinopyroxene) is also much lower in D-gabbro-norites than in N-gabbro-norites (Fig. 2).

From these observations, it can be concluded that there is a first-order difference in the composition of the parental liquids of the two cumulate suites. N-gabbro-norites present the classical character of evolved cumulates from tholeiitic melts, cogenetic with the troctolites and olivine gabbros found inside the diapir, although some samples with high Mg number (Mg/(Mg + Fe)) may require buffering processes related to wall rock assimilation⁹. The abundance of orthopyroxene in the D-gabbro-norites and its early appearance in the crystallization sequence, not related to the Fe (oxides) and water (hornblende) enrichment as in the N-suite, implies liquids with a higher SiO₂ content, more 'andesitic', than MORBs (typically ~55%). Also, the low Na and Ti concentration in minerals of the D-suite suggests that their parent melts issued from a source extremely depleted in incompatible elements.

In order to further characterize this source, we have performed trace element and Nd and Sr isotopic measurements on samples from the two suites of cumulates (Fig. 3). Values for ϵ_{Nd} (¹⁴³Nd/¹⁴⁴Nd normalized to 100-Myr-old chondritic reservoir) range from +7.52 to +9.41 for N-gabbro-norites and +6.56 to +9.03 for D-gabbro-norites (with the exception of sample 95 M

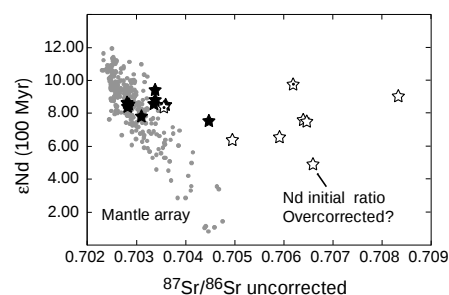


Figure 3 ϵ_{Nd} versus $^{87}\text{Sr}/^{86}\text{Sr}$ for the Maqsd cumulates. Symbols as in Fig. 1. ϵ_{Nd} of Maqsd cumulates show a narrow variation range characteristic of the oceanic mantle ($\epsilon_{\text{Nd}} = ^{143}\text{Nd}/^{144}\text{Nd}$ normalized to 100 Myr-old chondritic reservoir). $^{87}\text{Sr}/^{86}\text{Sr}$ shifts to the right of the mantle correlation, up to $^{87}\text{Sr}/^{86}\text{Sr}$ of 0.704 for olivine gabbros and clinopyroxene-troctolites. N-gabbro-norites show the same trend, while D-gabbro-norites-websterites plot in a distinct field ($^{87}\text{Sr}/^{86}\text{Sr}$ from 0.70357 to 0.70832). A shift in $^{87}\text{Sr}/^{86}\text{Sr}$ toward high values, falling outside the mantle correlation trend, is classically attributed to hydrothermal alteration by sea water ($^{87}\text{Sr}/^{86}\text{Sr}$ of Cretaceous sea water ranges from 0.706 to 0.709; refs 10, 11). In order to reduce the effect of post-crystallization alteration, all the analytical procedures, from the field selection to the final analyses, were conducted with extreme care. The samples in the field were selected for freshness and the final selection was based on thin-section examination. %LOI (loss on ignition) was also used as a freshness test. All rock chips (~5 mm in diameter) were dust-leached (bi-distilled H₂O plus ethanol) and hand-picked according to strict criteria: no carbonate veins, no secondary phases. ϵ_{Nd} were calculated assuming an age of 100 Myr for the ophiolite formation. Sr isotopic ratios were measured on leached clinopyroxene separates, and not age corrected (Rb/Sr ratios are very low). They were leached following the protocol defined by Snow *et al.*²⁵. Nd isotopic ratios were not measured on leached clinopyroxene, because leaching can change the Sm/Nd ratios and hence introduce errors into the corrected Nd ratios. Measured ratios are corrected for the laboratory bias determined from the La Jolla standard (present value $^{143}\text{Nd}/^{144}\text{Nd} = 0.511950$, Toulouse) for Nd and NBS 987 (present value $^{87}\text{Sr}/^{86}\text{Sr} = 0.71025$, Brest) for Sr. Blanks are 20 pg for Nd and 245 pg for Sr.

150A, for which the age correction may have been overestimated). These ranges fall within the range of other Maqsd cumulates ($+5.99 \leq \epsilon_{\text{Nd}} \leq +11.7$) and are typical of MORB values, especially Indian MORBs. This confirms a mantle origin for the two suites. In contrast with their similar Nd isotope values, the two gabbro-norite families have different Sr isotopic compositions. Sr isotopic ratios for N-gabbro-norites fall between 0.70282 and 0.70446, in the range of other Maqsd cumulates ($0.70291 \leq ^{87}\text{Sr}/^{86}\text{Sr} \leq 0.70479$), while D-gabbro-norites fall in a much more radiogenic range ($0.70357 \leq ^{87}\text{Sr}/^{86}\text{Sr} \leq 0.70832$). Given the range in ϵ_{Nd} , values of $^{87}\text{Sr}/^{86}\text{Sr}$ higher than about 0.7035 fall outside the mantle correlation trend and are thus not attributable to a pure mantle source. Such high $^{87}\text{Sr}/^{86}\text{Sr}$ values have been attributed (in Oman and in other ophiolites)^{4,10,11} to post-crystallization hydrothermal alteration. In the present study, particular care was taken when selecting and preparing the samples (see details in Fig. 3 legend), so it seems likely that the Sr isotopic signature of the D-gabbro-norites is a characteristic inherited from the parent melt. This conclusion is supported by the correlation between $^{87}\text{Sr}/^{86}\text{Sr}$ values and the purely igneous characteristics: modal composition (Fig. 4a), [Nd]/[Yb]N for equilibrium liquids, where N indicates normalization to chondritic values (Fig. 4b), and the anorthite content of plagioclase (Fig. 4c). In these diagrams, N- and D-gabbro-norites plot in clearly distinct ranges, the N-gabbro-norites within the olivine gabbro-troctolite field, the D-gabbro-norites and pegmatitic websterites defining their own field. All but one of the cumulates displaying high modal orthopyroxene (>20%) and all but one of the cumulates having very high An% plagioclase (>0.90) have high $^{87}\text{Sr}/^{86}\text{Sr}$ ratios (>0.70591), and vice versa.

Equilibrium liquids with D-gabbro-norites present the most

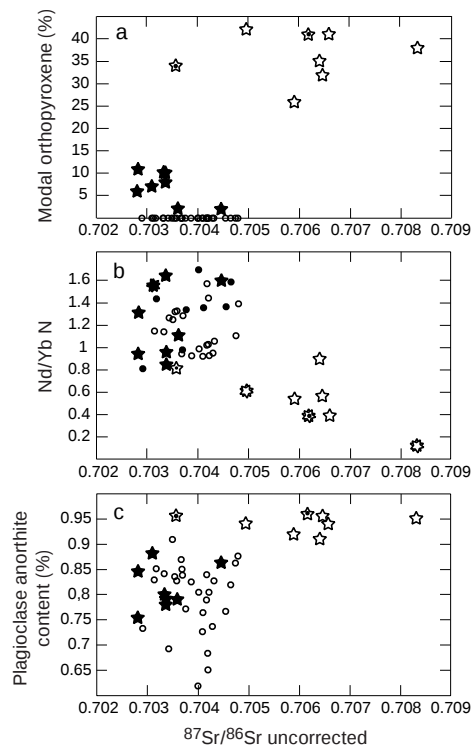


Figure 4 Orthopyroxene content (a), [Nd]/[Yb] N for equilibrium liquids (b), and anorthite content of plagioclase (c) versus $^{87}\text{Sr}/^{86}\text{Sr}$. The Nd and Yb concentrations are measured on whole rock (ICP-MS, Toulouse) and/or on clinopyroxene (10-branched stars and/or filled circles) (SIMS, WHOI, Woods Hole). Equilibrium liquids are calculated using partition coefficients^{20–24} and modal composition in the case of whole rock data. Normalization values are from ref. 26. Other symbols as in Fig. 1.

depleted signatures in terms of [Nd]/[Yb] N ratios, mostly associated with the high $^{87}\text{Sr}/^{86}\text{Sr}$ values (Fig. 4b). The ranges of [Yb] N values for N- and D-gabbro-norites overlap slightly (see Table 1) but the average values are clearly distinct. The sample presenting the lowest [Nd]/[Yb] N ratio is the one with the highest modal orthopyroxene content and the one with the most radiogenic $^{87}\text{Sr}/^{86}\text{Sr}$ ratio. Equilibrium liquids with N-gabbro-norites fall without exception in the field defined by liquids in equilibrium with Maqsd cumulates while equilibrium liquids with the D-gabbro-norites are strongly depleted in light rare earth elements (LREE), clearly distinct from cumulates in equilibrium with MORBs. A reverse correlation exists between the $^{87}\text{Sr}/^{86}\text{Sr}$ values and the [Nd]/[Yb] N ratios for the D-gabbro-norite and D-websterite samples.

Possible mechanisms for producing extremely depleted melts include (1) remelting the residue of a previous melting stage and (2) isolating the last (shallow) melt batches produced during decompression melting of a mantle diapir^{2,5}. The abnormally high $^{87}\text{Sr}/^{86}\text{Sr}$ values of Maqsd D-cumulates indicate the role of a fluid phase with a seawater signature during the genesis of these rocks and lead us to favour the first mechanism. Remelting of hydrated peridotites accounts for the SiO_2 enrichment required by early orthopyroxene crystallization, SiO_2 enrichment being a consequence of peridotite melting in the presence of water¹².

Now we consider the tectonic setting where this remelting event has taken place. It could be argued that Maqsd D-cumulates have the petrographical characteristics of arc cumulates. However, the very depleted LREE signatures of the melts in equilibrium with these cumulates contrast greatly with REE patterns observed in arc volcanism¹³. Although the role of a fluid phase seems inescapable, the details of the process are certainly quite different in Maqsd than in subduction zones. Moreover, in the case of late intrusion by arc-

related magmas, a random distribution of these intrusions relative to the Maqsd diapir would be expected, and the reverse is observed³. Finally, it is clear that an arc setting can be excluded for similar lithologies drilled at present-day ridges. The very depleted gabbro-norites drilled off the Mid-Atlantic Ridge (ODP site 334; ref. 2) are very similar to Maqsd D-gabbro-norites in terms of their mineralogy and major and trace element compositions. Also, highly anorthitic plagioclase megacrysts in MORBs are ubiquitous; their origin is still discussed¹⁴. Finally, SiO_2 -enriched and incompatible-element-depleted melts have been shown to exist at mid-ocean ridges as inclusions in olivine phenocrysts¹⁵.

Thus we propose that the Maqsd depleted suite formed in a pure mid-ocean ridge setting by remelting of a previously depleted lithosphere. The existence of thin axial lithospheric lids made of hydrated peridotites intruded by gabbroic plutons is increasingly documented along present-day slow-spreading ridges. They probably form during magma-starved stages separating two diapiric pulses^{16–18}. As an asthenospheric diapir reaches shallow depth, isotherms will rapidly rise in the overlying lithosphere and the hydrated peridotite solidus (hundreds of degrees colder than the dry solidus¹⁹) will eventually be crossed. Hydrated melting may affect the peridotites in a diffuse way. Such *in situ* 'lithospheric' melts have little chance of finding their way to the surface. Most of them are likely to remain trapped at sub-seafloor level where they crystallize as plutonic rocks and are eventually moved to an off-axis position as spreading progresses. This is consistent with the preferred occurrence of D-cumulates at the periphery of the diapir. The same process of hydrated melting can also occur in the vicinity of deep-seated faults after the diapir has cooled down.

While depleted melts are expected to remain largely isolated within the lithosphere, it is not unlikely that they are tapped from time to time by MORBs on their way to the surface. Mixing in various proportions between shallow depleted melts and pristine 'asthenospheric' melts may be proposed as a general mechanism that could contribute to the isotopic and geochemical variability of MORBs. However, reliable Sr isotopic data on depleted cumulates or melt inclusions sampled at present-day ridges are required before our hypothesis of hydrated melting at mid-ocean ridges can be confirmed. □

Received 23 April; accepted 5 October 1999.

- Dick, H. J. B., Robinson, P. T. & Meyer, P. S. The plutonic foundation of a slow-spreading ridge. *American Geophysical Union monograph* **70**, 1–39 (1992).
- Ross, K. & Elthon, D. Cumulates from strongly depleted mid-ocean-ridge basalt. *Nature* **365**, 826–829 (1993).
- Ceuleneer, G., Monnereau, M. & Amri, I. Thermal structure of a fossil mantle diapir inferred from the distribution of mafic cumulates. *Nature* **379**, 149–153 (1996).
- Benoit, M., Polvé, M. & Ceuleneer, G. Trace element and isotopic characterization of mafic cumulates in a fossil mantle diapir (Oman ophiolite). *Chem. Geol.* **134**, 199–214 (1996).
- Grove, T. L., Kinzler, R. J. & Bryan, W. B. in *Mantle Flow and Melt Generation at Mid-ocean Ridges* (eds Phipps Morgan, J., Blackman, D. K. & Sinton, J. M.) 281–310 (AGU Geophysical Monographs, 71, 1992).
- Ross, K. & Elthon, D. Cumulus and postcumulus crystallization in the oceanic crust: major- and trace-element geochemistry of Leg 153 gabbroic rocks. *ODP Sci. Res.* **153**, 333–350 (1997).
- Elthon, D. Mineral chemistry of gabbroic rocks from the Mid-Cayman Rise spreading center. *J. Geophys. Res.* **92**, 658–682 (1987).
- Amri, I., Benoit, M. & Ceuleneer, G. Tectonic setting for the genesis of oceanic plagiogranites: evidence from a paleo-spreading structure in the Oman ophiolite. *Earth Planet. Sci. Lett.* **139**, 177–194 (1996).
- Kelemen, P. Reaction between ultramafic rock and fractionating basaltic magma 1. Phase relations, the origin of calc-alkaline magma series, and the formation of discordant dunite. *J. Petrol.* **31**, 51–98 (1990).
- McCulloch, M. T., Gregory, R. T., Wasserburg, G. J. & Taylor, H. P. Jr. Sm-Nd, Rb-Sr, and $^{18}\text{O}/^{16}\text{O}$ isotopic systematics in an oceanic crustal section: evidence from the Samail ophiolite. *J. Geophys. Res.* **86**, 2721–2735 (1981).
- Noiret, G., Montigny, R. & Allègre, C. J. Is the Vourinos complex an island arc ophiolite? *Earth Planet. Sci. Lett.* **56**, 375–386 (1981).
- Dickson Spulber, S. & Rutherford, M. J. The origin of rhyolite and plagiogranite in oceanic crust: an experimental study. *J. Petrol.* **24**, 1–25 (1983).
- McCulloch, M. T. & Gamble, J. A. Geochemical and geodynamical constraints on subduction zone magmatism. *Earth Planet. Sci. Lett.* **102**, 358–374 (1991).
- Meyer, P. S. & Shibata, T. Complex zoning in plagioclase feldspars from ODP site 648. *ODP Sci. Res.* **106/109**, 123–142 (1990).
- Sobolev, A. V. & Shimizu, N. Ultra-depleted primary melt included in an olivine from the Mid-Atlantic Ridge. *Nature* **363**, 151–154 (1993).

16. Cannat, M. Emplacement of mantle rocks in the seafloor at mid-ocean ridges. *J. Geophys. Res.* **98**, 4163–4172 (1993).
17. Gente, P. *et al.* Characteristics and evolution of the segmentation of the Mid-Atlantic ridge between 20° N and 24° N during the last 10 Myr. *Earth Planet. Sci. Lett.* **129**, 55–71 (1995).
18. Linn, J., Purdy, G. M., Shouten, H., Sempere, J.-C. & Zervas, C. J. Evidence from gravity data for focused magmatic accretion along the mid-Atlantic ridge. *Nature* **344**, 627–632 (1990).
19. Willie, P. J. Magma genesis, plate tectonics and chemical differentiation of the Earth. *Rev. Geophys.* **26**, 370–404 (1988).
20. Irving, A. J. A review of experimental studies of crystal/liquid trace element partitioning. *Geochim. Cosmochim. Acta* **42**, 743–770 (1978).
21. Irving, A. J. & Frey, F. A. Trace element abundances in megacrysts and their host basalts: constraints on partition coefficients and megacryst genesis. *Geochim. Cosmochim. Acta* **48**, 1201–1221 (1984).
22. McKenzie, D. & O'Nions, R. K. Partial melt distributions from inversion of rare earth element concentrations. *J. Petrol.* **32**, 1021–1091 (1991).
23. Hart, S. T. & Dunn, T. Experimental cpx/melt partitioning of 24 trace elements. *Contrib. Mineral. Petrol.* **113**, 1–8 (1993).
24. Bédard, J. H. A procedure for calculating the equilibrium distribution of trace elements among the minerals of cumulate rocks, and the concentration of trace elements in the coexisting liquids. *Chem. Geol.* **118**, 143–153 (1994).
25. Snow, J. E., Hart, S. R. & Dick, H. J. B. Nd and Sr isotope evidence linking mid-ocean-ridge basalts and abyssal peridotites. *Nature* **371**, 57–60 (1994).
26. Nakamura, N. Determination of REE, Ba, Mg, Na and K in carbonaceous and ordinary chondrites. *Geochim. Cosmochim. Acta* **38**, 757–775 (1974).

Acknowledgements

Field work in Oman was conducted from 1992 to 1995, and was made possible thanks to facilities provided by the Ministry of Petroleum and Minerals of the Sultanate. We thank H. Al Azri for his constant support. Financial support was provided by the Centre National de la Recherche Scientifique in the frame of the “Dynamique et Bilan de la Terre” and “DORSALE” programs. We thank P. Kelemen for his invitation to perform ion probe analyses in Woods Hole. We thank P. Kelemen, B. Dupré, M. Monnereau and M. Rabinowicz for discussions. Many thanks also to the technical staff of Toulouse and Brest for help during data acquisition: M. Valladon (chemistry), P. Brunet and C. Bassoullet (mass spectrometry), B. Reynier (ICP-MS), P. de Parceval (microprobe), A.-M. Roquet (thin sections). We also thank M. Puel-Benoit for hand-picking fresh sample chips.

Correspondence and requests for materials should be addressed to G.C.
(e-mail: ceulener@gaia.cnes.fr).

An early Cambrian craniate-like chordate

Jun-Yuan Chen*, Di-Ying Huang* & Chia-Wei Li†

* Nanjing Institute of Palaeontology and Geology, Nanjing 210008, China

† Department of Life Science, National Tsing Hua University, Hsinchu, Taiwan, China

Since the identification of the Lower Cambrian *Yunnanozoon* as a chordate in 1995 (ref. 1), large numbers of complete specimens of soft-bodied chordates from the Lower Cambrian Maotianshan Shale in central Yunnan (southern China) have been recovered. Here we describe a recently discovered craniate-like chordate, *Haikouella lanceolata*, from 305 fossil specimens in Haikou near Kunming. This 530 million-year-old (Myr) fish-like animal resembles the contemporaneous *Yunnanozoon* from the Chengjiang fauna (about 35 km southeast of Haikou) in several anatomic features. But *Haikouella* also has several additional anatomic features: a heart, ventral and dorsal aorta, an anterior branchial arterial, gill filaments, a caudal projection, a neural cord with a relatively large brain, a head with possible lateral eyes, and a ventrally situated buccal cavity with short tentacles. These findings indicate that *Haikouella* probably represents a very early craniate-like chordate that lived near the beginning of the Cambrian period during the main burst of the Cambrian explosion. These findings will add to the debate on the evolutionary transition from invertebrate to vertebrate².

Genus *Haikouella* gen. nov.

Type species. *Haikouella lanceolata* gen. et sp. nov.

Etymology. The generic name refers to the fossil locality, Haikou.

Diagnosis. A small soft-bodied chordate a few centimetres long, which is lancelet-like, elongated and pointed at both ends (Fig. 1). The body is broadly triangular in its anterior half but narrowly triangular in the rest. It differentiates into a head, a trunk and a bent caudal projection. A distinct head bears possible lateral eyes and a ventral buccal cavity at its hind end. An expanded pharyngeal cavity bears six pairs of filamentous branchial arches. A large subventral notochord extends through the entire length of the trunk into a bent caudal projection. A circulatory system comprises ventral and dorsal aortae; a globular heart at the posterior end of the ventral aorta; and an anterior branchial arterial that runs in a circular course to connect the dorsal and ventral aorta at their anterior ends. The trunk bears a neural cord, which extends throughout the entire length of the trunk and into the head region, and forms an expanded brain anteriorly, which appears to be tripartite. The alimentary canal is differentiated into oesophagus, spiral midgut and a straight intestine. Four pairs of gonads are not metamerically arranged.

Remark. This new chordate shows a striking resemblance to the contemporaneous *Yunnanozoon*^{1,3}. Both have a large subventral notochord, massive dorsal musculature divided by nearly straight myosepta into about two dozen myomeres, and a large pharyngeal cavity with six or seven pairs of branchial arches. Unlike *Yunnanozoon*, however, this new chordate has fewer and non-metamerically arranged gonads, a wider ventral region in the anterior part of the body, smaller and more anteriorly situated pharyngeal teeth (Ba3) and distinctive gill filaments.

Haikouella lanceolata gen. et sp. nov.

Holotype. Complete specimen (Figs 2a and 3p), 27.5 mm long, oriented with sagittal plane subvertical to bedding in anterior but subhorizontal in the rest (EC00213a,b).

Material. A total of 305 specimens under study include: 30 complete specimens, 32 nearly complete, 110 representing the anterior part of the body (mainly pharynx, a few with the head region), 30 with only isolated preservation of branchial arches and ventral aorta, and 103 with only a part of the trunk.

Etymology. The specific name refers to lancelet-shaped animal (*L.*, lanceolatum).

Diagnosis. Same as in the generic diagnosis.

Locality and stratigraphy. Ercai Village, Haikou, Kunming; Lower Cambrian Maotianshan Shale.

Occurrence. Most *Haikouella* specimens are derived in a mud bed about 2.5 cm thick. Like most of the soft-bodied fossil-bearing layers in the Chengjiang fauna^{4,5}, it represents a microturbidite deposit, which comprises a lower, graded unit and an upper, homogeneous mud unit. An abundant occurrence of algal remains in the lower unit indicates the presence of a eutrophic environment in the burial site, whereas the *Haikouella* specimens at an aggregate occurrence in the upper mud unit were evidently transported by turbidite mudflow from an adjacent shallower area.

Description. Most adult specimens of *Haikouella lanceolata* are 25–30 mm long, although few reached a length of 40 mm. *Haikouella* differs from *Yunnanozoon* by having a broadly ventral region in its anterior part (with a dorsal angle of about 70°). The broadly triangular anterior part was buried either with its lateral side or its broad ventral region lying parallel to the plane of bedding, and thus compacted ventrolaterally or subdorsally. Behind the pharyngeal region, the trunk narrows to a dorsal angle of about 20° to 30°, usually compacted laterally. The transitional region of the two different parts of the trunk at the hind part of the pharynx is usually preserved as a twist (Figs 2a, c, g–j; 3a, b, e) as a result of *post mortem* compaction.

The dorsal compaction is extremely uncommon both in the

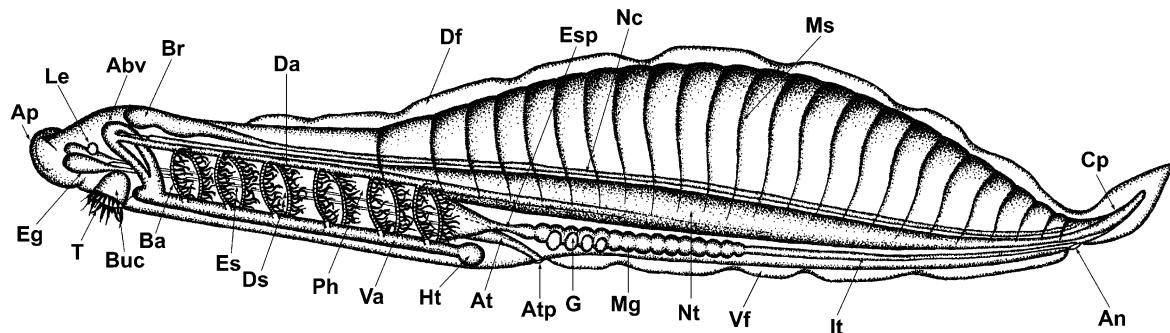


Figure 1 Anatomical interpretation of *Haikouella lanceolata* (gen. et sp. nov.) from Haikou, near Kunming. Abbreviations (also used in Figs 2–4): Abv, anterior branchial vessel; An, anus; Ap, anterior projection; At, atrio; Atp, atriopore; Ba, branchial arches; Baf, branchial-arch-filamental; Br, brain; Buc, buccal cavity; Co, copulatory organ; Cp, caudal project; Da, dorsal aorta; Df, dorsal fin; Ds, denticular structure; Eg, endostyle

glands; Es, endostyle; Esp, oesophagus; G, gonad; Hd, head; Ht, heart; It, intestine; Lb, lobated structures; Le, lateral eye; Mg, midgut; Mm, myomeres; Mo, mouth opening; Ms, myosepta; Mw, median wall; Nc, neural cord; Nt, notochord; Ph, pharyngeal cavity; T, tentacle-like structure; Va, ventral aorta; Vf, ventral fin.

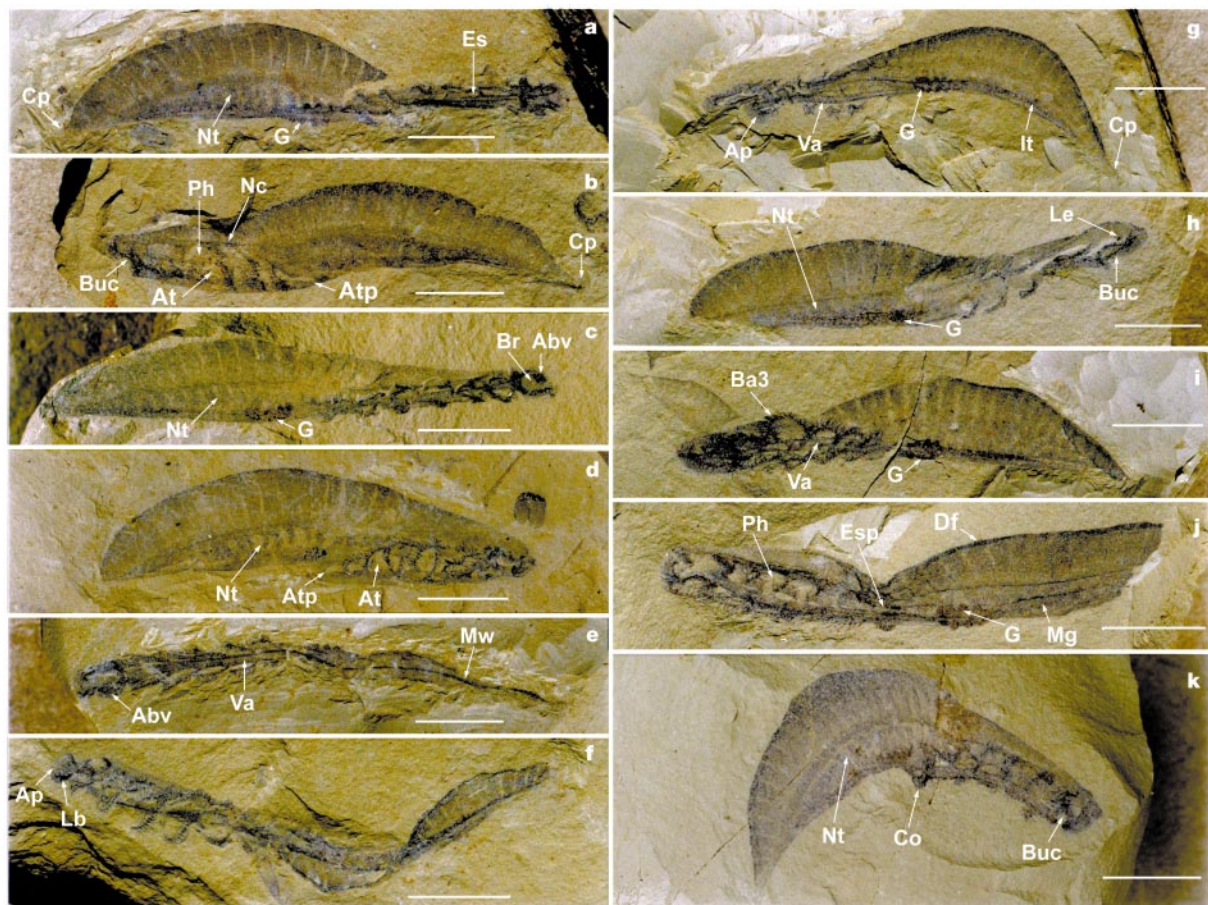


Figure 2 *Haikouella lanceolata* (gen. et sp. nov.) from Haikou, near Kunming. Abbreviations are as Fig. 1. Scale bars, 5 mm. **a**, Holotype (EC00213a), a sublateral specimen with subdorsally compressed anterior part, showing Ap, Lb, Br, Eg, Es, G, Nt, Cp and Ba. **b**, A sublateral specimen (EC00041a), showing mud-filled At and Ph, Atp, Nc, Cp and Buc. **c**, A sublateral specimen (EC0002a) but with the anterior part (before Ba6) compacted lateroventrally, showing Abv, Br, G and Nt. **d**, A sublateral specimen (EC00258a), showing mud-filled At, Atp, Nt and Ba. **e**, A dorsally compacted specimen

(EC00231), showing Ap, Va, Ba and Mw running internally through the trunk. **f**, A twisted specimen (EC00047a), showing Ap and Lb. **g**, A sublateral specimen (EC00048a) showing Ap, Va, G, It and Cp. **h**, A sublateral specimen (EC00001a), showing Buc, Nt and an eye on lateral side of head. **i**, A sublateral specimen with a subventral anterior part (EC00042a), showing Ba, Va, G and It. **j**, EC00118a compacted subventrally before Ba6 with the rest compacted sublaterally, with a twist between the two parts, showing Ph, Esp, Df, Vf, G and Mg. **k**, A ventrally curved specimen (EC00012a), showing Co, Buc and Nt.

present species (in less than 2%) and in *Yunnanozoon*. Figure 3o shows a ventral view of a dorsally compacted animal that is in a laterally sinuous posture, with traces of axial structures interpreted as notochord and intestine. The dorsally compacted animal (Fig. 3f), which is relatively straight, bore a dorsal fin compacted into a ridge-like structure running mid-dorsally along the trunk. The same ridge-like structure is also seen in a dorsally compacted specimen

of a possible *Yunnanozoon* in Chengjiang and interpreted as a subdorsal notochord. These dorsally compacted specimens, especially those in laterally sinuous preservation, may represent individuals that were killed by a sudden burial event. As in *Yunnanozoon*⁷, most of the *Haikouella* fossils are preserved mostly as bluish-grey flattened films.

The head region is well preserved in about 10% of the specimens,

either in dorsal (Figs 2a, e–g; 3p) or in sublateral compaction (Figs 2b, h; 3h, j, m). The head, as seen in ventral view, includes an anterior bulb-shaped region (referred as an anterior projection) that is set off from the rest of the head with paired lateral constrictions (Figs 2a, e–g; 3p). The ventral side of the anterior projection is divided into a pair of ventrolateral lobed structures (Figs 2a, e, f; 3p) that are similar to the bilobed anterior end of the conodont animal⁶. The ventral side of the head bears a dark-stained indentation (Figs 2b, h, j, k; 3g, h, j, m), referred to here as a buccal cavity,

that appears to be fringed with short tentacles (Fig. 3g, j) each about 0.6–0.7 mm long. The buccal cavity opens directly into the pharynx. A well defined circular structure (0.2 mm in diameter) is present on the lateral side of a head (Figs 2h and 3h), which we interpret as an indication of a lateral eye. In several specimens, the rock is split on a plane that passes through the body, revealing a fine black axial structure (about 0.1 mm in diameter), interpreted as a neural cord (Figs 2b; 3a, j), which was slender in the trunk. A large brain relative to body size is well represented by an anterior thickening of the

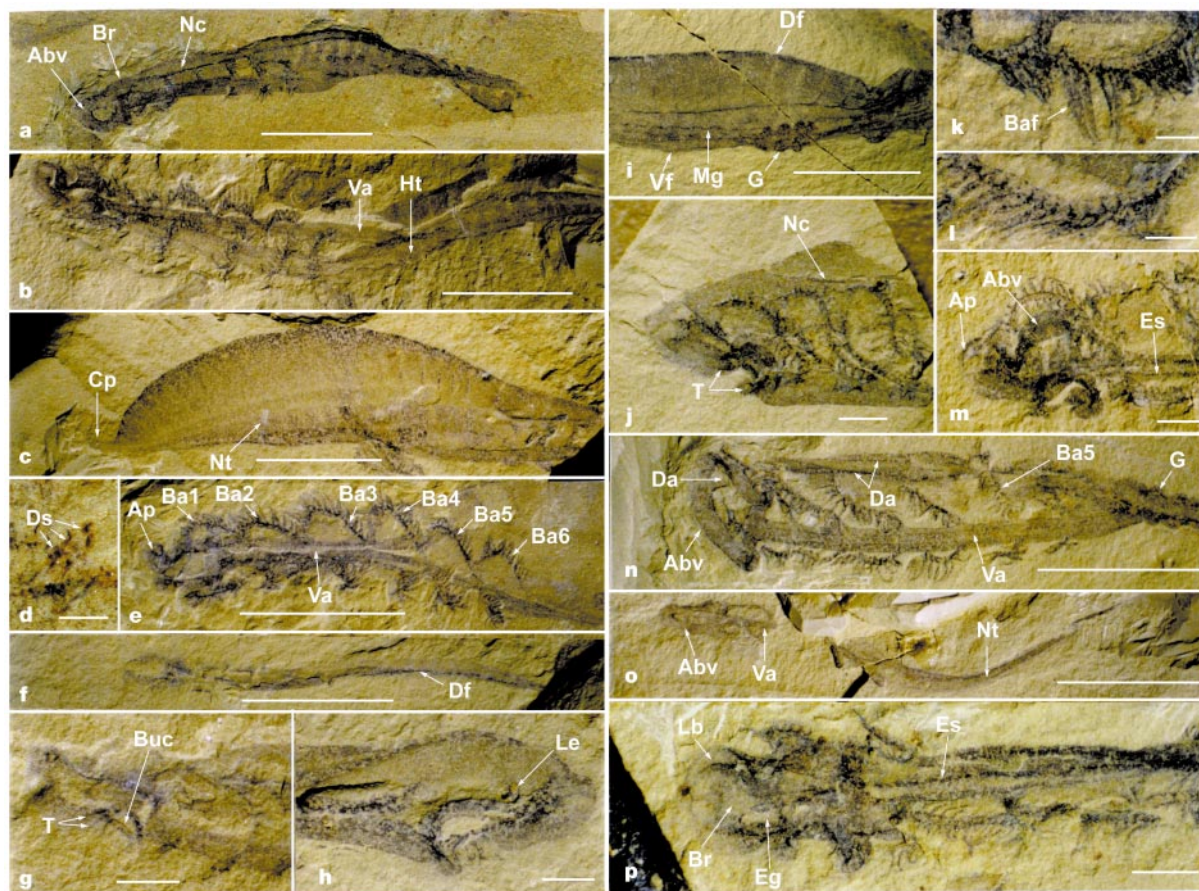


Figure 3 Views of *Haikouella lanceolata* (gen. et sp. nov.) from Haikou, near Kunming. Abbreviations are as in Fig. 1. Scale bars, 5 mm (a–c, e, f, i, n, o), 1 mm (g, h, j, m, p) and 0.5 mm (d, k, l). **a**, A twisted specimen (EC00214a), showing Abv, Br and Nc. **b**, EC00043a, showing Va and Ht. **c**, lateral specimen (EC00007a), showing Nt and a small Cp. **d**, Enlargement of Fig. 2b (EC00041a), showing Ds. **e**, Anterior part of EC00035, showing six pairs of filamentous branchial arches. **f**, A dorsal specimen (EC00031a), showing Df. **g**, Anterior part of EC00027, showing T and Buc. **h**, Enlargement

of Fig. 2h (EC00001a), showing one of lateral eye. **i**, EC00118b, counterpart of Fig. 2j showing G, Mg, Df and Vf. **j**, EC00300, showing T, Ba and Nc. **k**, EC00024a, showing Baf. **l**, Enlargement of a jointed Ba (EC00093b) that consists of narrow, dark disc-like bars and light wider spaces in between. **m**, EC00072a, showing Es and Ba with paired filaments. **n**, EC00157 showing Va, Da, Abv and Ba. **o**, Ventral view of a dorsally sinuous specimen (EC00050), showing Nt and Abv. **p**, Enlargement of the anterior part of EC00213a (Fig. 2a) showing Lb, Br, Eg and Es.

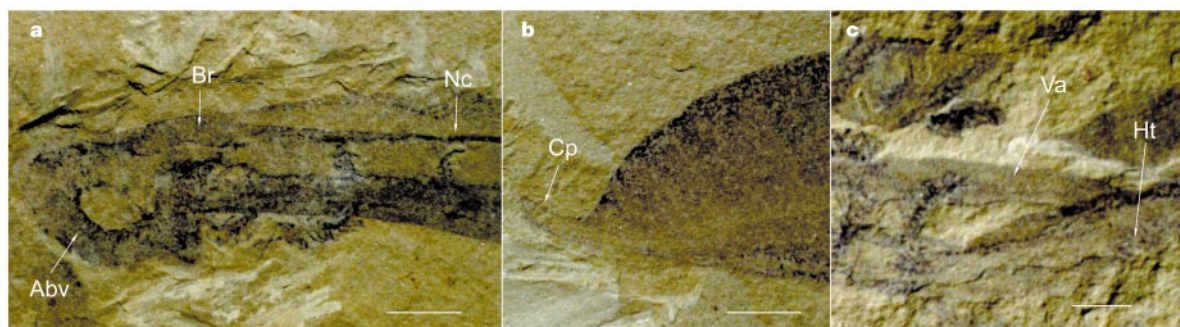


Figure 4 Enlarged view of *Haikouella laneolata* (gen. et sp. nov.) from Haikou, near Kunming. Abbreviations are as in Fig. 1. Scale bars, 1 mm. **a**, Enlargement of the anterior part of EC00214a (Fig. 3a), showing Br, Nc and Abv. **b**, Enlargement of the posterior end

of EC00007a (Fig. 3c), showing Cp. **c**, Enlargement of the posterior part of the pharyngeal cavity in EC00043a (Fig. 3b), showing Ht.

neural cord, and appears to tripartite (Fig. 4a). This elongated brain is about 4 mm long, and is differentiated longitudinally into three parts. The most posterior part is wedge-like, about 2 mm long, and lies at anterior part of the trunk. The anterior part of the brain is relatively thick, and is divided into two parts, both of which are lobate and 0.5 mm thick (Figs 2a, c; 3a, p; 4a).

As in *Yunnanozoon*⁷, the notochord of *Haikouella* is large and located relatively ventrally in the greater part of the trunk and runs into the caudal projection. The notochord is widest in the middle and tapers progressively towards its anterior and posterior ends. The caudal projection (Figs 2a, b, g, h, k; 3c; 4b) is a small but prominent structure, extending posteriorly, and is sometimes bent dorsally (Fig. 4b). The notochord is preserved as a flattened band which in most specimens is concealed under the muscle blocks that lie next to it (Figs 2a, c, d, h, k; 3c). As a result of the differentiated compaction of the musculature, an axial line may indicate the dorsal margin of the underlying notochord.

Fins are represented in only a few specimens. They comprise a narrow dorsal fin running along the entire length of the trunk and a narrow ventral fin running along the posterior half of the trunk, stopping at an anterior structure that may be an atriopore (Figs 2j; 3i).

An anterior mud-filled cavity is interpreted as a perforate pharynx (Fig. 2b, d, j) used for food collection and respiration. The pharynx occupies nearly the entire anterior half of the trunk before opening posteriorly into the short oesophagus (Fig. 2j). Six pairs of branchial arches (Figs 2a, c–f; 3a, b, e, n) slope forward slightly from either side of a pair of axial structures that may represent dorsal aortae (Fig. 3n). The ventral end of each arch attaches to either side of an axial structure regarded as a ventral aorta (Figs 2c, e, g, i–k; 3a, b, e, n, o). The branchial arches of *Haikouella* include a jointed structure (Fig. 3l) that resembles the striped gill-bar mucocartilage in ammocoete larvae⁸ and consists of about 25 disc-like dark bars with wider light spaces in between (Fig. 3l). The branchial arches are consistently well preserved. The ventral end of each bar is attached to the ventral aorta, but the dorsal end is readily detached and relocated (Figs 2f–h; 3a). On a given branchial arch, each of the dark discs bears two posteriorly projecting gill filaments about 1 mm long that are flattened dorsoventrally (Fig. 3k, m). In the floor of the pharynx is a possible endostyle, represented by a pair of ridges (Figs 2a; 3m, p), each with an expanded, teardrop-shaped anterior end (Fig. 3p). The pharyngeal cavity narrows to form a tube-like structure interpreted as an oesophagus (Fig. 2j) at its posterior end between B6 and the first gonad (G1). The

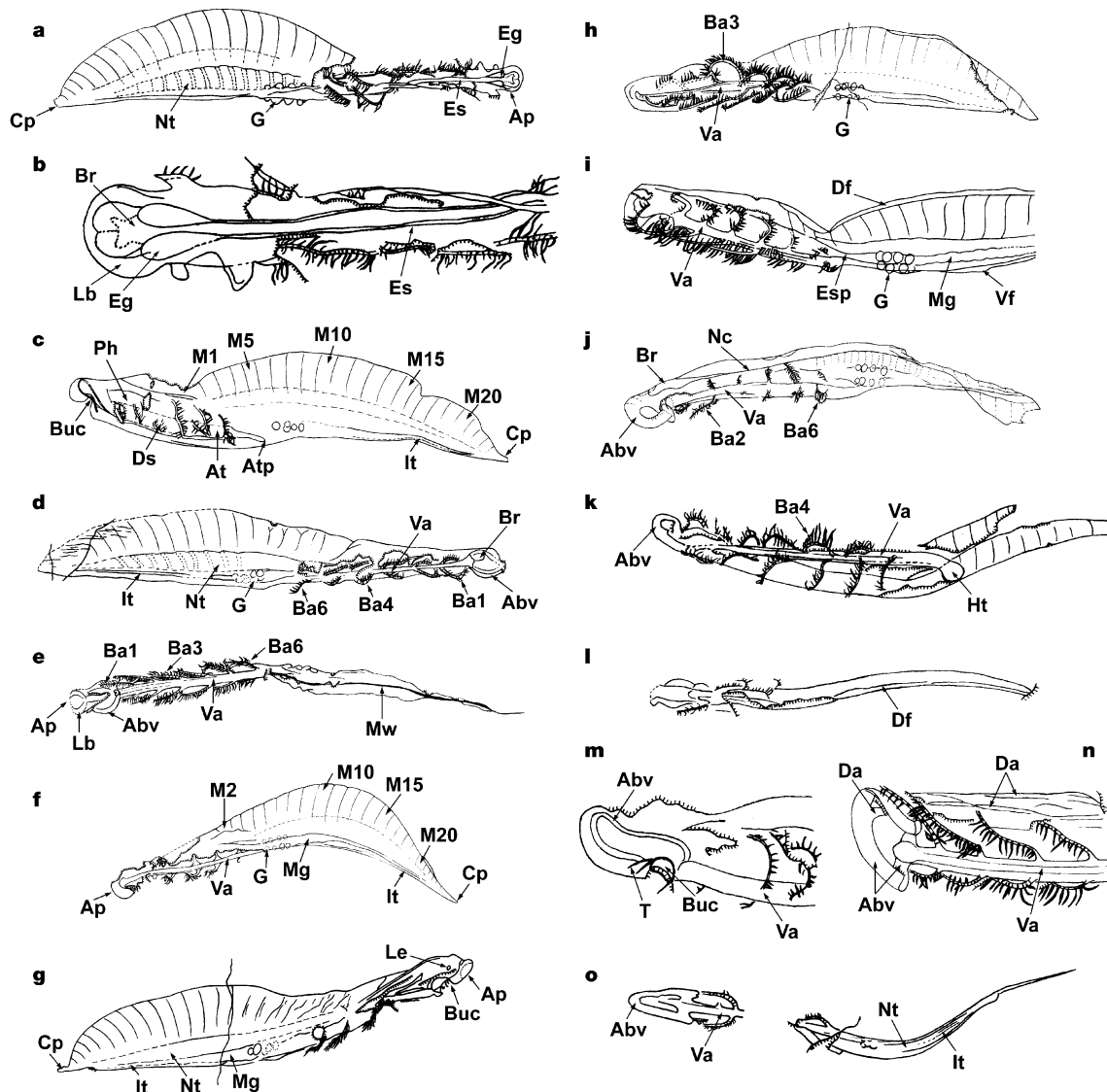


Figure 5 Camera lucida drawings of *Haikouella lanceolata* (gen. et sp. nov.) from Haikou, near Kunming. Diagrams are of the following figure parts: **a**, Fig. 2a; **b**, Fig. 3p; **c**, Fig. 2b;

d, Fig. 2c; **e**, Fig. 2e; **f**, Fig. 2g; **g**, Fig. 2h; **h**, Fig. 2i; **i**, Fig. 2j; **j**, Fig. 3a; **k**, Fig. 3b; **l**, Fig. 3f; **m**, Fig. 3g; **n**, Fig. 3n; **o**, Fig. 3o.

oesophagus leads directly to a spirally shaped midgut (Figs 2j and 3i) and a slender, straight intestine (Fig. 2b, c, g–i).

As in *Yunnanozoon*, there are simple cone-shaped denticular structures (Fig. 3d) within the pharynx, but those of *Haikouella* are much smaller, only 0.1 mm across and situated near B3 in a more anterior position within pharynx; in *Yunnanozoon* these structures are much larger, about 1 mm across, and situated at the posterior end of the pharyngeal cavity. We regard them as pharyngeal teeth. Midgut and intestine both lie immediately below the notochord, extending along its ventral margin to the ventral anus near to the posterior end of the trunk. The gill slits between the gill arches, instead of being exposed externally, were covered by lateral folds of the body (Figs 2k and 3c) which enclose a ventral atrial space (Fig. 2b, d). This atrium was an elongated cavity, usually mud-filled, extending ventrally with a posterior opening, the atriopore at the level of M6 (Fig. 2b, d). In *Yunnanozoon*, we also interpreted the gill slits as opening into an atrial cavity, although they have also been interpreted as opening directly to the exterior⁹. Below the notochord are four paired gonads that are arranged non-metamerically, at the posterior end of the atrium, beneath M5 to M6 (Figs 2a, c, g–j; 3a, i, n). In contrast, *Yunnanozoon* has 13 pairs of gonads, arranged metamerically with the first pair located below M6 (ref. 2).

The blood circulatory system consists of a set of closed vessels. At the posterior end of the pharyngeal region is a globular posterior enlargement of the ventral aorta, representing the heart (Figs 3b and 4c). From the ventral aorta, the blood presumably ascended through the branchial arches via the branchial arteries to the paired dorsal aortae (Fig. 3n). Each of the branchial arches has an expanded ventral end (Figs 2g and 3n), which may be equivalent to the branchial heart of lancelets¹⁰. At the anterior ends of the ventral and dorsal aortae, an anterior branchial artery runs forward and appears to connect them (Figs 2c, e; 3a, b, g, m–o) along the dorsal margin of the buccal cavity into the dorsal margin.

The trunk musculature, which is divided by relatively straight myosepta into about 25 myomeres (Fig. 2b, d, g, k), is situated mainly above the notochord, but also extends ventrally on either side of the notochord (Figs 2a, c, d, h, k; 3c). In contrast, the traces of segmented musculature on the flanks of notochord in *Yunnanozoon* have been reported as a gut valve structure¹¹. Apart from the thin but broadly spaced M1 and M2, the remaining muscle segments are evenly spaced. As in *Yunnanozoon*, the septa are weakly sigmoidal, curving backwards dorsally and forwards ventrally. The musculature may be divided into two lateral sets by a median wall. A dorsally compacted animal is shown in Fig. 2a with a darker, axial line running internally through the trunk at a mid-dorsal position, which we interpret as an indication of the median wall.

There is a large papilla in front of the atriopore of a specimen (Fig. 2k). The rare occurrence of this structure indicates that it might be a possible copulatory organ, as in lamprey¹², that is temporarily developed during the mating season.

Although *Yunnanozoon* has been interpreted as a hemichordate¹¹, we believe that this hypothesis is incorrect⁷. For instance, the proposed gut contents may be a ventral view of the gonads, and the observed spiral structure is a trace of musculature next to the notochord⁷, not a part of the gut. There is little doubt that the rod in both *Haikouella* and *Yunnanozoon* is a notochord, as it lies dorsally to the pharyngeal cavity and intestine canal. It seems certain that *Haikouella* and *Yunnanozoon* are closely related and that both are indeed chordates. *Yunnanozoon* was previously classified as a cephalochordate¹, but this was challenged by its interpretation either as a basal chordate¹³ or as an independent class of the chordates⁹. Anatomic findings such as the brain and possible lateral eyes in its close relative *Haikouella* pose a question as to whether *Yunnanozoon* may also be an early craniate.

Although living chordates display an amazing diversity of body forms, extant lancelets are broadly accepted to be the best available proxies for the latest common ancestor of the cephalochordates and

craniates¹⁴. Palaeontological information on *Yunnanozoon* and *Haikouella*, as well as molecular genetic data¹⁵, have lent strong support to this concept. It is possible that all the craniate body forms evolved from a common ancestor resembling these fish-like species. Although it is commonly considered that the origin of the craniates was signalled by the relatively simultaneous appearance of a conspicuous brain and an endoskeleton including a cranium, our finding indicates that the craniates originated through a set of separate events over a long interval of time. The appearance of a conspicuous brain may have been the earliest of these events, occurring long before full endoskeletization. Although the lancelet has sometimes been regarded as a brainless chordate, recent micro-anatomical and molecular-genetic studies indicate that lancelets (and, by extension, the proximate common ancestor of the cephalochordates and craniates) have at least a diencephalic forebrain¹⁶ and a relatively large hindbrain¹⁷. *Haikouella* had a relatively large brain and possibly a pair of lateral eyes, which indicate that this animal might be considered as an early craniate.

The origin of the craniates has been broadly accepted as being linked to defence in such forms as the armoured heterostracans and may have evolved in the late Cambrian and flourished in the Silurian–Devonian^{10,12}. However, the first hard parts of the craniates appear to have been linked to the feeding function. The eel-like, soft-bodied conodonts, extended back into the late Cambrian by fossil remains of the euconodont elements, may represent one of the early craniates that were able to build tooth-like calcium phosphate in the mouth, allowing them to adopt to an active hunting life¹⁸. The pharyngeal teeth in the pharyngeal cavity of both *Yunnanozoon* and *Haikouella* may represent the earliest known biomineralization in the chordate. □

Received 10 June; accepted 24 September 1999.

- Chen, J. Y., Dzik, J., Edgecombe, G. D., Ramsköld, L. & Zhou, G.-Q. A possible Early Cambrian chordate. *Nature* **377**, 720–722 (1995).
- Holland, N. D. in *International Symposium of the Origins of Animal Body Plans and their Fossil Records* (eds Chen, J.-Y. et al.) 11 (Early Life Research Center, Jinning, 1999).
- Hou, X.-G., Ramsköld, L. & Bergström, J. Composition and preservation of the Chengjiang Fauna—Lower Cambrian soft-bodied biota. *Zool. Scripta* **20**, 395–411 (1991).
- Chen, J.-Y., Zhou, G.-Q., Zhu, M.-Y. & Yeh, G.-Y. *The Chengjiang Biota—A Unique Window of the Cambrian Explosion* (National Museum of Natural Sciences Press, Taichung, 1996).
- Chen, J.-Y. & Zhou, G.-Q. in *The Cambrian Explosion and the Fossil Record* (eds Chen, J.-Y., Cheng, Y.-N. & Van Iten, H.; *Bull. Natl. Mus. Nat. Sci.* **10**, 11–105, (1997).
- Aldridge, R. J. & Purnell, M. A. The conodont controversies. *Trends* **11**, 463–466 (1996).
- Chen, J.-Y. & Li, C.-W. in *The Cambrian Explosion and the Fossil Record* (eds Chen, J.-Y., Cheng, Y.-N. & Van Iten, H.; *Bull. Natl. Mus. Nat. Sci.* **10**, 257–273, 1997).
- Dohrn, A. Studien zur Urgeschichte des Wirbelthierkörpers. V. Zur Entstehung und Differenzierung der Visceralbogen bei Petromyzon planeri. *Mitt. Zool. Station Neapel* **5**, 152–161 (1884).
- Dzik, J. *Yunnanozoon* and the ancestry of chordates. *Acta Palaeontol. Pol.* **40**, 341–360 (1995).
- Blik, A. At the origin of chordates. *Geobios* **25**, 101–113 (1992).
- Shu, D.-G., Zhang, X.-L. & Chen, L. Reinterpretation of *Yunnanozoon* as the earliest known hemichordate. *Nature* **380**, 428–429 (1996).
- Young, J. Z. *The Life of Vertebrates* 3rd edn (Clarendon, Oxford, 1981).
- Chen, J.-Y. in *International Cambrian Explosion Symposium* (eds Chen, J.-Y., Edgecombe, G. & Ramsköld, L.) 7–9 (Early Life Research Center, Nanjing, 1995).
- Stokes, M. D. & Holland, N. D. The lancelet. *Am. Sci.* **86**, 552–560 (1998).
- Wada, H. & Satoh, N. Details of the evolutionary history from invertebrates to vertebrates, as deduced from the sequence of 18S rDNA. *Proc. Natl Acad. Sci. USA* **91**, 383–490 (1994).
- Locall, T. C., Holland, N. D. & West, J. E. Landmarks in the anterior central nervous system of amphioxus larvae. *Phil. Trans. R. Soc. Lond. B* **344**, 165–185 (1994).
- Holland, P. W. H., Holland, L. Z. & Holland, N. D. The molecular control of spatial patterning in amphioxus. *J. Mar. Biol. Assoc. UK* **74**, 49–60 (1994).
- Purnell, M. A., Aldridge, R. J., Donoghue, P. C. J. & Gabbott, S. E. Conodonts and the first vertebrates. *Endeavour* **19**, 20–27 (1995).

Acknowledgements

We thank N. Holland, E. Davidson, and D. Walossek for discussions and comments, and F. Gao for technical assistance. This work was supported by National Department of Science and Technology, National Foundation of Natural Sciences, Jiangsu Provincial Committee of Science and Technology. The research of L.C.W. is supported by National Science Council.

Correspondence and requests for material should be addressed to C.J.Y. (e-mail: chenjy@jlonline.com). All the fossil material has been deposited at the Early Life Research Center, Jinning.

Sympatric speciation by sexual selection

M. Higashi, G. Takimoto & N. Yamamura

Center for Ecological Research, Kyoto University, Kyoto 606-8502, Japan

There is increasing evidence^{1–6} for the process of sympatric speciation^{7,8}, in which reproductive isolation of species occurs without physical isolation. Theoretical models^{9–14} have focused on disruptive natural selection as the crucial pressure for splitting a species. Here we report the theoretical finding that sympatric speciation may be caused by sexual selection even without disruptive natural selection. Specifically, we show that variation in a

male secondary sexual character with two conspicuous extremes and the corresponding variance in female mating preference around no preference may jointly evolve into bimodal distributions with increasing modal divergence of the male and female traits, pulling a population apart into two prezygotically isolated populations. This mode of speciation, driven by two runaway processes^{15–17} in different directions, is promoted by an increase in the efficiency of females in discriminating among males or a decrease in the cost of male conspicuousness, indicating that sympatric speciation may occur more readily if barrier-free or predator-free conditions arise. Although even a slight cost of female preference would cancel the runaway process of sexual selection¹⁸, it would not cancel the divergent runaway processes of sympatric speciation.

Consider a diploid species with polygynous mating (that is, each female chooses one mate and breeds once, but the entire set of males

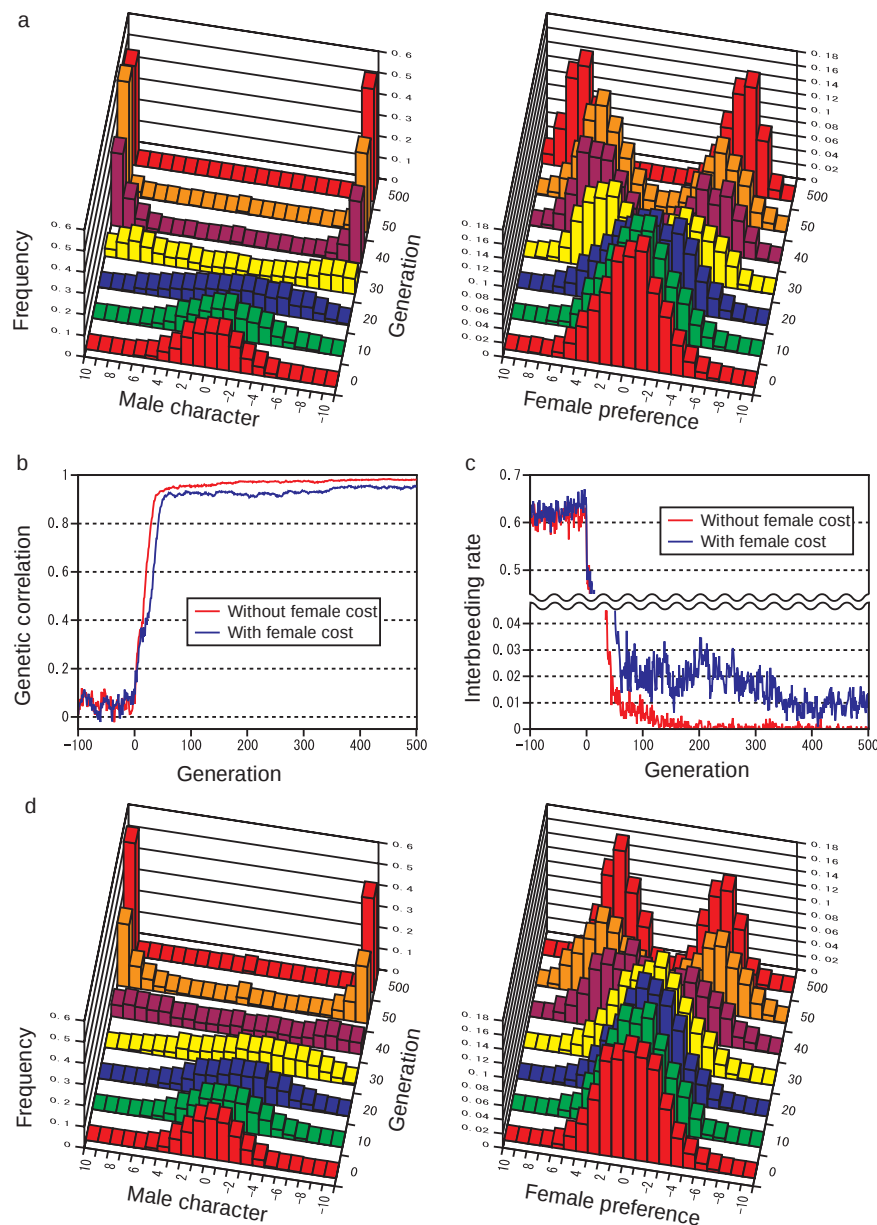


Figure 1 Examples of sympatric speciation generated by our model. We set $N = 1,000$ and $m = n = 5$. The initial distributions of male and female mating traits (x and y) were generated as their stationary distributions with $\alpha = 0.02$ and $\omega_M = 10$; the evolutionary dynamics of speciation were generated with $\alpha = 0.2$ and $\omega_M = 10$. **a**, The evolutionary changes in the distributions of male and female mating traits in the case without female

cost. **b**, **c**, The genetic correlation between male and female mating trait loci (**b**) and the evolutionary change in the interbreeding rate between the two subpopulations, each of which comprises the males and females of the same sign at the breeding season (**c**), in cases without and with female cost ($\omega_F = 20$). **d**, The evolutionary changes in the distributions of male and female mating traits in a case with female cost ($\omega_F = 20$).

is available to every female). Suppose¹⁹ that males have a secondary sexual character x with two extremes, whose value is additively determined by m loci (each taking three alleles: $-1, 0, +1$) as $-2m \leq x \leq 2m$ (where $-$ and $+$ represent alternative directions of the male trait), and that females have the corresponding alternative mating preferences, y , additively determined by n loci (each taking three alleles: $-1, 0, +1$) as $-2n \leq y \leq 2n$. For example²⁰, in some fish males range from black to white through different shades of grey; the extremely black and white males, represented by $-2m$ and $+2m$, respectively, in our model, are the most conspicuous and suffer the highest predation. Females have varying degrees of mating preferences for either black or white males; the most choosy females for black and white males are represented in our model by $-2n$ and $+2n$, respectively, and $y = 0$ means no preference.

Assume that for one-time mating a female of mating preference y chooses a male of trait value x out of the entire set of surviving male adults with a probability proportional to $e^{\alpha xy}$ (that is, the probability $e^{\alpha xy} / \sum_x p(x) e^{\alpha xy}$, where $p(x)$ denotes the number of males of trait value x at the breeding season, which varies from generation to generation, and α represents the efficiency of male discrimination by females)¹⁶. She produces a fixed number of offspring, each of which is assigned a sex with a 1 : 1 ratio and receives from each of its parents one genome randomly chosen after free recombination.

We assume a cost to the male of his secondary sexual character x , which is reflected in the relative premating survival probability (under predation pressure) of a male of trait value x , $\exp[-x^2/2\omega_M^2]$, where ω_M denotes the range of x with high viability¹⁶. We assume that there is no cost to the female of her mating preferences¹⁸ (although we will later consider such a cost). N male and N female adults are randomly chosen as survivors at the breeding season out of the male and female offspring populations, respectively, weighted by their relative survival probability distributions.

Although we assume here that there are no mutations (mutations from ± 1 to 0 and their back mutations), numerical simulations show that the following theoretical results are also true with mutations. Also, the adoption of an alternative mate choice by females according to the 'best of n rule'¹⁷ (that is, a female courts with n males and chooses the one with the greatest xy value) would lead to the same main conclusions, where the male discrimination efficiency of females is reflected in the sample size n . (In fact, speciation occurs more readily with the best of n rule as long as n is sufficiently large, because discriminating females avoid males of the opposite sign more strongly when conspicuous males appear.)

Starting with a pair of unimodal distributions around $x = y = 0$ for the male and female mating traits (x and y), generated as the stationary distributions for a case in which speciation does not occur, our model simulates the evolutionary process of speciation under certain conditions. In such a case, the initial unimodal distributions are transformed into bimodal ones whose modal distance increases with time (Fig. 1a), where the genetic correlation (linkage disequilibrium) between the loci of male secondary sexual character and those of female mating preference builds up quickly (Fig. 1b). Also, the interbreeding rate between the two subpopulations, each of which comprises the males and females of the same sign at the breeding season (that is, the ratio of the frequency with which females mate with a male of the opposite sign to that of all mating events) rapidly declines almost to zero, indicating that prezygotic isolation is quickly established (Fig. 1c).

The population bifurcation is driven by divergent runaway processes. There is mutual reinforcement of male and female mating traits such that more conspicuous males increase because they are preferred by discriminating females of the same sign (while intermediate males decrease), and then more discriminating females increase by hitchhiking (owing to increasing linkage disequilibrium). This in turn enhances the advantage of more conspicuous males over intermediate males.

With the same initial unimodal distributions around $x = y = 0$, three distinct evolutionary outcomes may occur: no change, trait shift (in a stochastically determined direction) and speciation. Which occurs will depend on the change in the values of parameters involved (Fig. 2). Given the other parameter values, speciation may occur if the efficiency α of male discrimination by females exceeds a threshold. As α increases from 0 to the threshold the frequency of trait shift increases relative to that of no change; after α exceeds the threshold speciation occurs with increasing frequency and then replaces trait shift (Fig. 3a). Speciation may be prevented by too high a cost of male secondary sexual character (that is, too small a value of ω_M), and be promoted by a reduction in male cost (Fig. 3b). (Note that the likelihood of speciation is greater in the case of Fig. 3a than in that of Fig. 3b for the same parameter values $\alpha = 0.2$ and $\omega_M = 10$. This is because the variances of the initial distributions are greater in the former case (in which the initial distributions are generated under $\alpha = 0.02$ and $\omega_M = 10$) than in the latter case (in which the initial distributions are generated under $\alpha = 0.2$ and $\omega_M = 1$), indicating that greater genetic variability in the original population should facilitate speciation.)

These dependencies of the model dynamics on the parameter values suggest the following evolutionary mechanism: a species population with moderate mating traits is sustained as long as the male discrimination efficiency of females is sufficiently low, or the predation pressure on conspicuous males is sufficiently high. If these barriers are weakened (owing to a change in the physical or biological environment) the population may either shift its mating traits, if the change is limited, or undergo speciation. Numerical simulations show that if a trait shift occurs following a change in either of these parameters, a further change in that parameter (in the same direction) will not lead to speciation (although it is possible to return to the original state by reversing the parameter change and then produce speciation by a large change in either of the parameters, noting that the evolutionary processes of trait shift and speciation are reversible, as indicated in Fig. 2). Thus, speciation may occur with rapid, not gradual, changes in these key parameters.

A larger population size, N (more strictly, more males or females participating in reproduction), would facilitate speciation (Fig. 3c) because it would reduce the random drift in the trait distributions. Furthermore, numerical simulations (performed with scaling for comparison) confirm that a change in the remaining parameters m and n , the numbers of the loci determining the male and female mating traits, respectively, would not affect the evolutionary outcome.

It might be thought that the mechanism for sympatric speciation

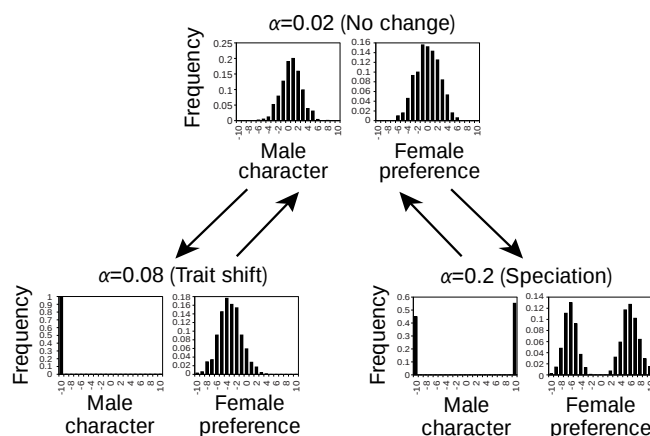


Figure 2 Three possible evolutionary outcomes: no change, trait shift and speciation. The evolutionary transitions (indicated by arrows) may occur with changes in the parameters α as indicated, while fixing the other parameters as $\omega_M = 10$, $N = 1,000$ and $m = n = 5$.

driven by divergent runaway processes hinges on the symmetry of the dipolar female mating preference, which is the ultimate force for pulling the population apart in two directions, or on that of male cost, which in contrast suppresses the growth of divergent runaway processes. Speciation may, however, occur even if the female mating preferences or male costs on the – and + sides of the trait ranges are not exactly balanced: that is, $\alpha^- \neq \alpha^+$ or $\omega_M^- \neq \omega_M^+$, where α^- and α^+ (ω_M^- and ω_M^+) are, respectively, the α (ω_M) values for the negative and positive y (x) values (Fig. 3d, e). This indicates that this mechanism for sympatric speciation does not depend on this subtle balance, and that, even if the forces in the opposite directions are off balance, speciation may be achieved as long as prezygotic isolation can be established before either of the two subpopulations (on the negative and positive sides of the mating trait ranges) becomes extinct.

It is known¹⁸ that the incorporation of even a slight cost of female mating preference into a model for sexual selection by a runaway

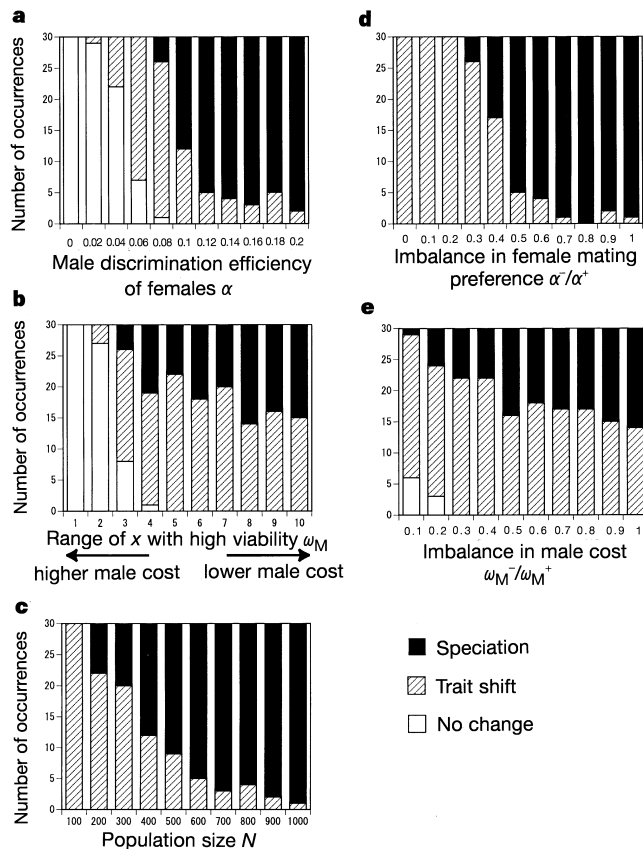


Figure 3 The likelihood (represented here by the number of occurrences out of 30 simulations runs) of each of the three evolutionary outcomes, speciation (defined here as a case in which the interbreeding rate is less than 0.05), trait shift and no change, when either α or ω_M is changed from a value for generating initial stationary unimodal distributions to another value. $m = n = 5$. **a**, The results when α is changed from 0.02 (for generating the initial distributions) to the values indicated while fixing $\omega_M = 10$, $N = 1,000$. **b**, The results when ω_M is changed from 1 (for generating the initial distributions) to the values indicated while fixing $\alpha = 0.2$, $N = 1,000$. **c**, The results when α is changed from 0.02 (for generating the initial distributions) to 0.2, while fixing $\omega_M = 10$, under the different population sizes indicated. **d**, The results when α^- and α^+ are changed from 0.02 (for generating the initial distributions) to 0.2 and the values indicated on the horizontal axis multiplied by 0.2, respectively, while fixing $\omega_M = 10$, $N = 1,000$. α^- and α^+ , respectively, denote the α value for the negative y values and that for the positive y values. **e**, The results when ω_M^- and ω_M^+ are changed from 1 (for generating the initial distributions) to 10 and the values indicated on the horizontal axis multiplied by 10, respectively, while fixing $\alpha = 0.2$, $N = 1,000$, where ω_M^- and ω_M^+ , respectively, denote the ω_M value for the negative y values and that for the positive y values.

process causes the system to return to its original state, cancelling the runaway process, and that additional factors have to be considered to prevent this cancellation^{21,22}. To examine the effect of this cost to the female, we incorporated it into our model for sympatric speciation in a similar way to male cost; that is, in terms of the relative survival probability of a female of value y , $\exp[-y^2/(2\omega_F^2)]$, where ω_F indicates the range of y with high viability. As long as it is sufficiently small (that is, ω_F is sufficiently large), a female cost would not cancel the divergent runaway processes, but only reduce the modal divergence of the male and female mating traits in the case of sympatric speciation by sexual selection (Fig. 1c, d). The reason for this difference in the effect of female cost between the case of a single runaway process and that of divergent runaway processes is inferred to be as follows: in the former case females can without any loss, avoid the costly mating preference when the male trait of the entire population evolves to its steady state (that is, the balance between sexual selection benefit from, and natural selection cost of the male mating trait). In the latter case they have to pay a greater cost for disadvantageous interbreeding (because of the presence of another peak on the opposite-signed side of the male mating trait range) as they reduce their mating preference intensity.

We have derived from a theoretical model a novel mechanism for sympatric speciation driven by divergent runaway processes. Our main conclusions also follow from a simplified deterministic model (G.T., M.H. and N.Y., in preparation), suggesting their robustness.

Wu²³ presented a sexual selection model which assumes that females have an absolute mating preference (that is, a female most prefers males with a male mating trait value uniquely specified by her preference trait value) and showed that sympatric speciation was possible. However, prezygotic isolation could not be sustained in the long term in a small population (in contrast to our model, in which relative mating preference is assumed), indicating that relative mating preference is required for lasting sympatric speciation.

Another model²⁰, which is similar to, but different in several key features from, our model, showed that sympatric speciation may occur if all females initially possess an extreme mating preference and the corresponding male secondary sexual character has a distribution biased toward the extreme preferred by the females. Assuming such a specific initial condition, this model leads to the conclusion that speciation requires a cost of male secondary sexual character; the authors considered this male cost to be the driving force for speciation (although it is not clear whether their conclusion, that is, their interpretation of the model results, is correct because they did not examine the effect of male cost by comparing cases with different male costs, including the case with no cost). Their conclusion is completely opposite to the prediction derived from our model that male cost is not necessary for, but its reduction promotes, speciation (Fig. 3b). This prediction leads to a hypothesis that sympatric speciation by sexual selection tends to occur as predation pressure is reduced (with other conditions unchanged) by, for example, a population entering into a new enemy-free habitat such as an island or a lake.

Another prediction derived from our model, that sympatric speciation is promoted by the efficiency of male discrimination by females (Fig. 3a), explains high species diversity in a habitat with good visibility (for example, cichlid fish diversity in the African great lakes^{2,3,6}), and implies that high species diversity depends critically on the integrity of the environment³. □

Received 2 September; accepted 14 September 1999.

1. Turner, G. F. Speciation mechanism in Lake Malawi cichlids: A critical review. *Arch. Hydrobiol.* **44**, 139–160 (1994).
2. Deutsch, J. C. Colour diversification in Malawi cichlids: evidence for adaptation, reinforcement or sexual selection? *Biol. J. Linn. Soc.* **62**, 1–14 (1997).
3. Seehausen, O., Alphen, J. J. M. & Witte, F. Cichlid fish diversity threatened by eutrophication that curbs sexual selection. *Science* **277**, 1808–1811 (1997).
4. Seehausen, O. & Alphen, J. J. M. The effect of male coloration on female mate choice in closely related Lake Victoria cichlids (*Haplochromis nyererei* complex). *Behav. Ecol. Sociobiol.* **42**, 1–8 (1998).

5. Ritchie, M. G. & Phillips, S. D. F. in *Endless Forms: Species and Speciation* (eds Howard, D. J. & Berlocher, S. H.) 291–308 (Oxford Univ. Press, New York, 1998).
6. Albertson, R. C., Markert, J. A., Danley, P. D. & Kocher, T. D. Phylogeny of a rapidly evolving clade: The cichlid fishes of Lake Malawi, East Africa. *Proc. Natl Acad. Sci. USA* **96**, 5107–5110 (1999).
7. Tauber, C. A. & Tauber, M. J. in *Speciation and its Consequences* (eds Otte, D. & Endler, J. A.) 307–344 (Sinauer, Sunderland, 1989).
8. Bush, G. L. Sympatric speciation in animals: new wine in old bottles. *Trends Ecol. Evol.* **9**, 285–288 (1994).
9. Maynard Smith, J. Sympatric speciation. *Am. Nat.* **100**, 637–650 (1966).
10. Tauber, G. A. & Tauber, M. J. A genetic model for sympatric speciation through habitat diversification and seasonal isolation. *Nature* **268**, 702–705 (1977).
11. Doebeli, M. A quantitative genetic competition model for sympatric speciation. *J. Evol. Biol.* **9**, 893–909 (1996).
12. Kawecki, T. J. Sympatric speciation via habitat specialization driven by deleterious mutations. *Evolution* **51**, 1751–1763 (1997).
13. Kondrashov, A. S. & Kondrashov, F. A. Interactions among quantitative traits in the course of sympatric speciation. *Nature* **400**, 351–354 (1999).
14. Dieckmann, U. & Doebeli, M. On the origin of species by sympatric speciation. *Nature* **400**, 354–357 (1999).
15. Fisher, R. A. *The Genetical Theory of Natural Selection* (Clarendon, Oxford, 1930).
16. Lande, R. Models of speciation by sexual selection on polygenic traits. *Proc. Natl Acad. Sci. USA* **78**, 3721–3725 (1981).
17. Seger, J. Unifying genetic models for the evolution of female choice. *Evolution* **39**, 1185–1193 (1985).
18. Pomiankowski, A. The costs of choice in sexual selection. *J. Theor. Biol.* **128**, 195–218 (1987).
19. Liou, L. W. & Price, T. D. Speciation by reinforcement of premating isolation. *Evolution* **48**, 1451–1459 (1994).
20. Turner, G. F. & Burrows, M. T. A model of sympatric speciation by sexual selection. *Proc. R. Soc. Lond. B* **260**, 287–292 (1995).
21. Pomiankowski, A., Iwasa, Y. & Nee, S. The evolution of costly mate preferences I. Fisher and biased mutation. *Evolution* **45**, 1422–1430 (1991).
22. Iwasa, Y., Pomiankowski, A. & Nee, S. The evolution of costly mate preferences II. The 'handicap' principle. *Evolution* **45**, 1431–1442 (1991).
23. Wu, G.-I. A stochastic simulation study on speciation by sexual selection. *Evolution* **39**, 66–82 (1985).

Acknowledgements

M.H. thanks J. Lawton and the NERC Centre for Population Biology for their hospitality. This work was supported by a MESSC grant to M.H.

Correspondence and requests for materials should be addressed to M.H.
(e-mail: higashi@ecology.kyoto-u.ac.jp).

Direct measurement of intra-cochlear pressure waves

Elizabeth S. Olson

Physics Department, Princeton University, Princeton, New Jersey 08544, USA

The cochlear travelling wave is fundamental to the ability of the mammalian auditory system to resolve frequency. The seashell-shaped outer bone of the cochlea (the auditory inner ear) contains a spiral of cochlear fluid and the sensory tissue known as the cochlear partition. Sound travels down the ear canal to the eardrum, causing its flexible tympanic membrane to vibrate. This vibration is transmitted to the cochlea via the ossicles. Motion of the stapes (the stirrup ossicle) sets the cochlear fluid in motion, which in turn sets the cochlear partition near the stapes in motion. The motion of the cochlear partition ripples down the cochlear spiral as a travelling wave, stimulating the cochlea's sensory hair cells. The wave peaks near the base (the stapes end) of the cochlea for high frequency tones and near the apex for low frequencies¹. The fundamental elements of the cochlear travelling wave are fluid pressure and motion and partition forces and motion. However, the wave's direct experimental study has to date relied almost solely on measurements of the partition motion. Here I report finely spaced measurements of intracochlear pressure close to the partition, which reveal the fluid component of the cochlear wave. The penetration depth of the wave is very limited, $\sim 15\ \mu\text{m}$. Over a range of frequencies at least an octave wide, the depth is independent of frequency.

I recently described pressure measurements made in the extreme base of the cochlea²; here a turn-one location, 14% further towards the apex along the cochlear spiral, is emphasized. Strong nonlinearity was detected at the turn-one location, whereas in the extreme base the responses were nearly linear, owing to the extreme fragility of that region. Therefore, the basal results describe the passive cochlea. Nonlinearity in the cochlea's mechanical response to sound is the key to its dynamic range and ability to finely separate frequencies. Outer hair cells appear to be essential to this mechanical enhancement^{3,4}, whereas the inner hair cells communicate information about the partition's motion to the auditory neurons⁵.

Here I demonstrate the frequency and position dependence of the nonlinear fluid pressure. Pressure differences are used to find fluid velocity, and the results are interpreted with respect to the cochlear travelling wave. In particular, the wave's penetration depth is found to be $\sim 15\ \mu\text{m}$. The penetration depth indicates how much fluid the wave contains, and its size and frequency dependence constrain cochlear models. In models, the fluid component of the cochlea is often simplified to one dimension—along the partition. In two-dimensional models the height of the fluid is included, and in three-

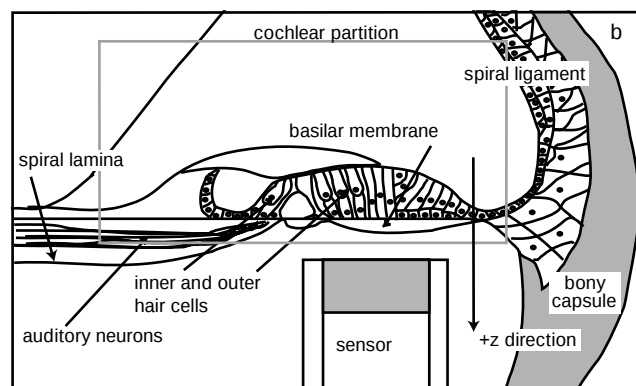
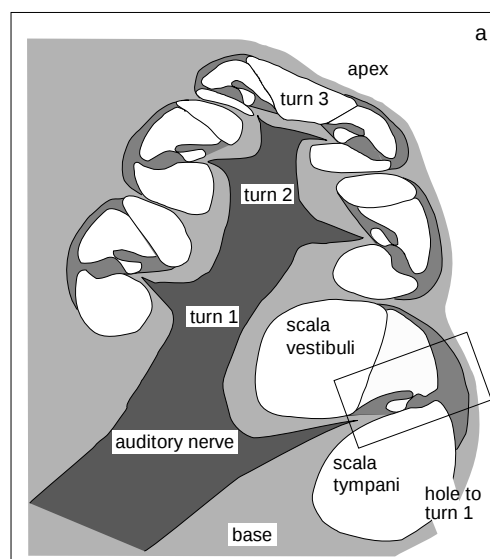


Figure 1 The cochlea. **a**, Cross-section of the gerbil cochlea; the boxed section is enlarged in **b**. The cochlear partition and the fluid compartments scala tympani and scala vestibuli spiral around the auditory nerve. The mongolian gerbil cochlea is $\sim 2.6\ \text{mm}$ in diameter, with a cochlear partition length of $\sim 13\ \text{mm}$ (ref. 16). In the extreme base of the cochlea the scala vestibuli is terminated by the stapes, the scala tympani by the round window membrane (not shown). The general location of a hand-drilled hole used to access turn one scala tympani is indicated in **a**. The scala tympani pressure sensor was positioned to approach the basilar membrane approximately perpendicularly. The sensor's outer diameter is $170\ \mu\text{m}$. (The drawings derive from a 2-mm plastic-embedded section prepared by M. Feldman.)

dimensional models the finite widths of the fluid compartments and cochlear partition are also used. The small, frequency-insensitive penetration depth emphasizes the significance of all three dimensions for modelling and understanding cochlear mechanics.

The measurements were made in deeply anaesthetized gerbils stimulated with tones via a speaker coupled to the ear canal. In each experiment one location along the cochlear spiral was probed in the scala tympani, either the extreme base or the first turn of the spiral (Fig. 1a). These basal and turn-one locations had characteristic frequencies of approximately 32 kHz and 20 kHz, consistent with Müller's place-frequency map of the gerbil cochlea⁶. The scala-tympani pressure measurements were made at a series of distances from the basilar membrane, a fibrous tissue that forms a boundary of the cochlear partition (Fig. 1b). The pressure in the scala vestibuli just next to the stapes was measured within seconds of each measurement in the scala tympani. This is the input pressure to the cochlea and served as a reference pressure. The results of one basal experiment and one turn-one experiment are shown in Figs 2–5, with the details of the basal experiment presented elsewhere².

Figure 2 shows turn-one scala-tympani pressure at the frequency of the stimulus, measured as the sensor approached the basilar membrane in 13 steps. For clarity, only five of the 13 curves are shown. The gain of the scala-tympani pressure relative to the pressure in the ear canal was as large as 50 dB. (For comparison, numerous experiments indicate that the gain of the scala-vestibuli pressure next to the stapes is ~30 dB and is almost flat at frequencies of 2–40 kHz (refs 2, 7).) At positions close to the basilar membrane the scala-tympani pressure peaked sharply at ~19 kHz when the stimulus level was 50 dB SPL, and peaked broadly in the frequency band of 10–20 kHz when the stimulus level was 80 dB SPL. (50–60 dB SPL is a conversational sound level; 80 dB SPL is the level of an alarm clock two feet away.) The difference in the sharpness of tuning occurred because at frequencies in the vicinity of the characteristic frequency the response magnitude did not scale linearly with stimulus level, but was compressed. This type of cochlear nonlinearity is well known from measurements of basilar-membrane motion⁸. The nonlinearity depends on physiological processes, because after death the pressure scaled linearly with stimulus level. Close to the basilar membrane, the phase of the scala-tympani pressure relative to the scala-vestibuli pressure decreased smoothly through more than a complete cycle as the frequency varied from 10 to 23 kHz. This accumulating phase is expected for a travelling wave, and is similar to the phase of basilar-membrane motion⁹. The notches that appeared in the magnitude at

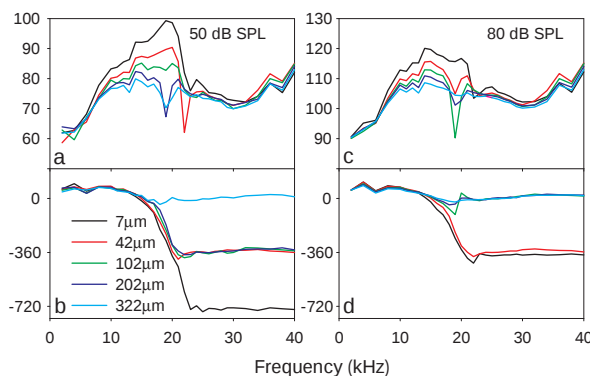


Figure 2 Scala tympani pressure versus frequency at the turn-one location. Each curve was measured at the distance from the b.m. indicated in the key. Pressure was also measured at 22, 62, 82, 122, 142, 162, 182, and 302 μm from the basilar membrane. **a, b**, Magnitude and phase, 50 dB SPL stimuli (as measured in ear canal); **c** and **d**, magnitude and phase, 80 dB SPL stimuli. Magnitude is expressed as SPL, defined as dB with respect to 20 μPa . The phase is referenced to the scala vestibuli pressure next to the stapes. (Scala tympani pressure with an accumulating phase such as that illustrated here has been measured in five basal and 13 turn-one experiments.)

frequencies in the vicinity of the characteristic frequency are believed to be caused by destructive interference between two modes of cochlear pressure². One of the modes is coupled to the cochlear travelling wave. The phase of this mode goes through large excursions (referenced to the input pressure in the scala vestibuli) and its magnitude decreases with distance from the basilar membrane. The other mode is more closely tied to the stapes motion and its magnitude and phase are expected to vary in space relatively slowly¹⁰. In this interpretation, the substantial pressure at frequencies well above the characteristic frequency, where the phase accumulation ceased, was due to the non-travelling wave mode. The pressure peak was substantially smaller and the phase no longer accumulated smoothly at distances 100–300 μm from the basilar membrane. Spatial pressure variations indicate fluid motion, and the rapid decrease of pressure with distance demonstrates the localized nature of the travelling wave disturbance in the fluid. At frequencies above the characteristic frequency, although the pressure was large, pressure variations were very small—substantial fluid motions were only present with the travelling wave. To quantify the fluid motions, the velocity of the fluid in the direction perpendicular to the basilar membrane was found from the spatial pressure variations. Fluid velocity is important because the rapidity with which it falls off with distance from the basilar membrane gives the penetration depth of the travelling wave in the fluid. The direct use of pressure to find the penetration depth of the travelling wave is confounded by the large non-travelling pressure mode.

The calculation of the z -component of fluid velocity begins by taking the difference between two pressures from adjacent positions along the z -axis (Fig. 1b); for example, those measured 7 μm and 22 μm from the basilar membrane. This difference divided by the distance between the measurements approximates the z -component of the pressure gradient. The gradient is related to fluid velocity via the Navier-Stokes equation

$$\nabla P = -\rho \partial v / \partial t + -\rho v \cdot \nabla v + \mu \nabla^2 v \quad (1)$$

where P is pressure, v is velocity, ∇ is the gradient operator, ∇^2 is the laplacian operator, and ρ and μ are the density and viscosity of the cochlear fluid. At frequencies above ~8 kHz, ∇P can be approximated using only the first term on the right-hand of this equation, and the velocity can be written very simply². Here P_b and P_a are two pressures measured at positions adjacent to each other on the z -axis with P_a further from the basilar membrane, and v_z is the z -component of the velocity. The difference Δz is the distance between the pressure measurements. The time derivative is written as $i\omega$ where ω is angular frequency

$$v_z = (i/\omega)(P_a - P_b)/(\rho \Delta z) \quad (2)$$

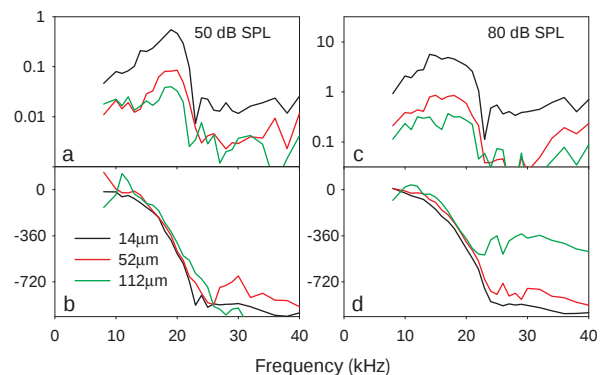


Figure 3 Fluid velocity derived from pressures of Fig. 2. Magnitude and phase of the derived velocity (z -component) are shown as a function of frequency at three distances from the basilar membrane. The phase is referenced to the scala vestibuli pressure next to the stapes. **a, b**, Magnitude and phase, 50 dB SPL stimuli; **c, d**, magnitude and phase, 80 dB SPL stimuli.

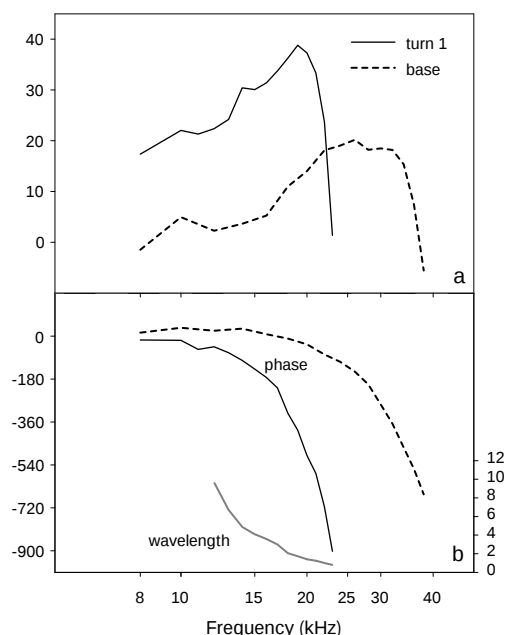


Figure 4 Fluid velocity (z -component) close to the basilar membrane at turn-one and basal locations, and wavelength of cochlear wave. **a**, **b**, Magnitude and phase. The magnitude is normalized to the stimulus level in the ear canal. The phase is referenced to the scala vestibuli pressure next to the stapes. Stimulus levels were 50 dB SPL (turn one) and 60 dB SPL (base). Inset in **b**, and using the right axis, is the wavelength of the cochlear travelling wave as a function of frequency. The derived wavelength corresponds to a location on the cochlear spiral that is midway between the basal and turn-one locations.

Figure 3 shows v_z as a function of frequency and distance from the basilar membrane, calculated with pairs of pressure measurements from the experiment of Fig. 2. The velocity magnitude did not contain the frequency and distance dependent notches which were seen in the pressure magnitude and were attributed to destructive interference between travelling-wave and non-travelling-wave pressure modes. The reason for the contrast is simple. Fluid velocity is based on the pressure gradient: spatial pressure differences. Only the travelling-wave mode (rapidly varying spatially) contributes substantially to the gradient; the non-travelling mode (slowly varying spatially) cancels out the difference. This cancellation also caused the velocity phases to go through larger and smoother excursions compared to their pressure counterparts.

I also found the wavelength and penetration depth of the travelling wave. Figure 4 shows the fluid velocity close to the basilar membrane for the turn-one experiment of Figs 2 and 3, and for a basal experiment. The fluid very close to the basilar membrane is expected to move with it. Thus, these results estimate basilar-membrane velocity. The validity of this estimation is supported by the similarity between these results and those of direct measurements of basilar-membrane velocity from the literature^{11,12}. It is possible to estimate the travelling wave wavelength from the phase data and the observation that the two locations (basal and turn one) were spaced by ~ 1.8 mm. Then $\lambda = 1.8(360/\Delta\phi)$, where λ is the wavelength in mm and $\Delta\phi$ is the phase difference between the two locations in degrees. This wavelength belongs to a location on the cochlear spiral that is midway between the basal and turn one locations, where the characteristic frequency would be ~ 25 kHz. Wavelength is plotted against frequency in Fig. 4. The wavelength decreased by a factor of ten as the frequency roughly doubled. This rapidly decreasing wavelength reflects the decreasing speed of the travelling wave as the stimulus frequency approached the characteristic frequency¹⁰. At 23 kHz, the wavelength is found to be just under 1 mm, which is similar to previously reported wavelengths

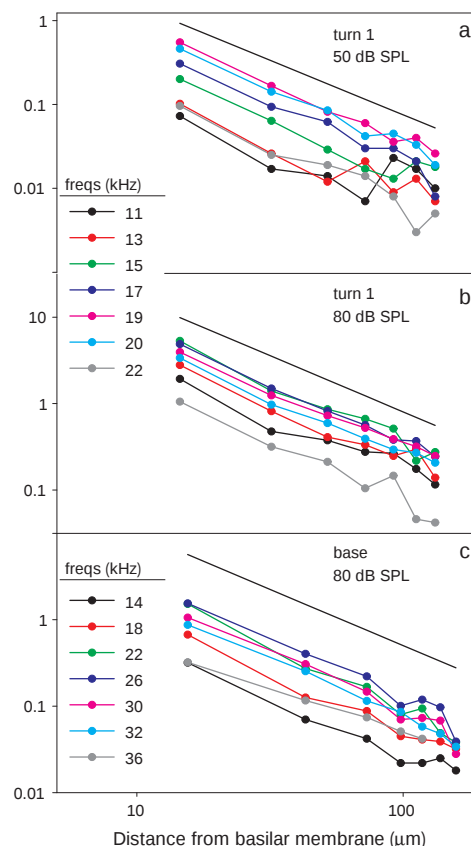


Figure 5 Fluid velocity (z -component) as a function of distance from the basilar membrane. **a**, Turn-one location, 50 dB SPL stimuli. **b**, Turn-one location, 80 dB SPL stimuli. **c**, Basal location, 80 dB SPL stimuli. The black line without data points is a power law with exponent -1.3 . (The fall-off of fluid velocity with distance illustrated here has been repeated in the four additional turn-one experiments that were part of this study.)

near the characteristic frequency in the basal region of squirrel monkey⁸, cat and guinea pig¹³. (The fact that the basal pressure behaved linearly is not expected to change the wavelength prediction greatly, because the response phase is relatively insensitive to stimulus level and cochlear condition (Fig. 3 and ref. 9).)

Figure 5 shows the magnitude of the velocity versus distance from the basilar membrane at frequencies from almost an octave below the characteristic frequency to just above it. The reduction in velocity with distance was smooth and rapid. It was quite insensitive to frequency, stimulus level and location along the spiral. (Less than 30 μ m from the basilar membrane, the fluid velocity increased with distance in some experiments. These short-range disturbances are the subject of further study.) Included in the velocity-distance graphs is a line representing a power law with exponent -1.3 . The penetration depth, δ , is defined as the distance from the closest velocity measurement to the place where the velocity is reduced by a factor of e . From the power law, $\delta = 15$ μ m, a distance less than one-tenth the width of the cochlear partition and smaller than the length of a hair cell.

The insensitivity of the penetration depth to frequency is remarkable, particularly when interpreted with respect to cochlear models. As in an ocean wave, the penetration depth is expected to depend on wavelength and, as Fig. 4 makes clear, the wavelength is sensitive to frequency. However, the width of the partition, ~ 200 μ m, is smaller than even the smallest wavelength. The results here suggest that it is this dimension that governs the penetration depth, and intuitively, it is reasonable that the smaller dimension should rule. More rigorously, a quantity from cochlear modelling that is related to the penetration depth is the 'effective mass' of the fluid. In one- and

two-dimensional models, the effective mass falls steeply as wavelength decreases. In three-dimensional models, in a fairly broad frequency range near the response peak, the effective mass becomes substantially less sensitive to wavelength^{14,15}. In light of this behaviour, and the intuitive argument above, the observed lack of frequency sensitivity in the penetration depth does not seem all that surprising. □

Methods

Pressure sensor

A sensor consists of a glass capillary (inner and outer diameters 100 and 170 μm , respectively) tipped with a gold-coated polymer diaphragm. Light from an LED is delivered via a fibre optic threaded into the capillary, and reflects from the diaphragm. The amount of light returning to the fibre optic for transmission to a photodetector varies linearly with the pressure-induced bending of the diaphragm. The acoustic impedance of the sensors is an order of magnitude larger than that of the cochlea, and their presence does not 'load' the cochlea outright. Nevertheless, in some animals small reversible changes in compound action potential (CAP) threshold and/or scala vestibuli pressure occurred when the scala tympani sensor was close to the basilar membrane. However, the derived fluid velocity close to the basilar membrane was similar to that measured directly by others^{11,12}, suggesting that the sensor's presence does not cause large changes in cochlear mechanics.

Experimental procedure

Animal procedures were approved by the Princeton University IACUC. The experimental animals were young adult gerbils. A gerbil was deeply anaesthetized and its left cochlea was exposed. Tones from a loudspeaker were delivered to the ear via a tube fitted to the left ear canal. The level of the tones was calibrated in the ear canal at the beginning of each experiment. Basal scala tympani pressure measurements were made by inserting a pressure sensor through the round window opening after removing the covering membrane. Scala vestibuli and turn-one scala tympani measurements were made through small holes hand-drilled in the cochlear bone. In basal experiments it was possible to see the basilar membrane in order to position the scala tympani sensor. In turn-one experiments it was positioned by using anatomical landmarks and referring to widely opened excised cochleae. It was not practical to systematically check the precision of positioning in each experiment. However, the grouped data indicate that incorrectly positioning the sensor towards the spiral ligament caused substantial damage, and that incorrectly positioning the sensor over the spiral lamina caused greatly diminished pressure gradients. The distance between the basilar membrane and the sensor was determined by touching the former with the latter, which produced a characteristically noisy signal.

Stimulus generation and recording was performed with a Tucker Davis Technologies DA/AD system. With typical signal averaging times of 3 s, sound pressure above the level of 60–70 dB SPL (20–60 mPa) could be reliably measured.

As a gauge of cochlear health, an electrode at the round window measured the CAP response of the auditory nerve to tones. The CAP threshold is the minimum sound level required to elicit a reliable neural response. Initial thresholds in the turn one experiment were ≤ 30 dB SPL at 15 and 20 kHz, and 50 dB SPL at 25 kHz. Close to the time of measurements, these thresholds were 40, 40, and 50–60 dB SPL. Initial thresholds in the basal experiment were ≤ 60 dB SPL at 30 kHz, and 80 dB SPL at 40 kHz. Close to the time of measurements, they were 60 and > 80 dB SPL. At the frequencies of interest, the initial CAP thresholds in the turn one experiment were in keeping with those of Müller⁶, those in the basal experiment were elevated somewhat. The degree of nonlinearity observed in the scala tympani pressures was consistent with the health of the cochlea as indicated by the CAP threshold—the basal experiment was nearly linear, and the nonlinearity in the turn-one experiment was strong, but not among the strongest in the literature⁹.

Received 1 June; accepted 4 October 1999.

1. von Békésy, G. *Experiments in Hearing* (McGraw Hill, New York, 1960).
2. Olson, E. S. Observing middle and inner ear mechanics with novel intracochlear pressure sensors. *J. Acoust. Soc. Am.* **103**, 3445–3463 (1998).
3. Brownell, W. E., Bader, C. R., Bertrand, D. & de Ribaupierre, Y. Evoked mechanical responses of isolated cochlear outer hair cells. *Science* **227**, 195–196 (1985).
4. Liberman, M. C. & Dodds, L. W. Single neuron labeling and chronic cochlear pathology. III. Stereocilia damage and alterations of threshold tuning curves. *Hear. Res.* **16**, 55–74 (1984).
5. Spoendlin, H. Innervation densities of the cochlea. *Acta Otolaryngol. (Stockh.)* **73**, 235–248 (1972).
6. Müller, M. The cochlear place-frequency map of the adult and developing mongolian gerbil. *Hear. Res.* **94**, 148–156 (1996).
7. Dancer, A. & Franke, R. Intracochlear sound pressure measurements in guinea pigs. *Hear. Res.* **2**, 191–205 (1980).
8. Rhode, W. S. Observations of the vibration of the basilar membrane in squirrel monkeys using the Mossbauer Technique. *J. Acoust. Soc. Am.* **49**, 1218–1231 (1971).
9. Ruggero, M., Rich, N. C., Recio, A., Narayan, S. S. & Robles, L. Basilar membrane responses to tones at the base of the chinchilla cochlea. *J. Acoust. Soc. Am.* **101**, 2151–2163 (1997).
10. Lighthill, J. Energy flow in the cochlea. *J. Fluid Mech.* **106**, 149–213 (1981).
11. Sellick, P. M., Yates, G. K. & Patuzzi, R. The influence of Mossbauer source size and position on phase and amplitude measurements of the guinea pig basilar membrane. *Hear. Res.* **10**, 101–108 (1983).
12. Xue, S., Mountain, D. C. & Hubbard, A. E. Electrically evoked basilar membrane motion. *J. Acoust. Soc. Am.* **97**, 3030–3041 (1995).
13. Cooper, N. P. & Rhode, W. S. Basilar membrane mechanics in the hook region of cat and guinea-pig

cochlea: Sharp tuning and nonlinearity in the absence of baseline position shifts. *Hear. Res.* **63**, 163–190 (1992).

14. de Boer, E. Auditory physics. Physical principles in hearing theory. II. *Phys. Rep.* **105**, 141–226 (1984).
15. Steele, C. R. & Taber, L. A. Comparison of WKB calculations and experimental results for three-dimensional cochlear models. *J. Acoust. Soc. Am.* **65**, 1007–1018 (1979).
16. Plassman, W., Peetz, W. & Schmidt, M. The cochlea in gerbilline rodents. *Brain Behav. Evol.* **30**, 82–101 (1987).

Acknowledgements

I thank H. Nakajima, E. de Boer, L. Sohn, N. Cooper, S. Staggs, R. Brawer, R. Austin and L. Page Jr. This work was supported by the National Institute on Deafness and Other Communication Disorders.

Correspondence and requests for materials should be addressed to E.O. (e-mail: colson@princeton.edu).

Stable propagation of synchronous spiking in cortical neural networks

Markus Diesmann*, Marc-Oliver Gewaltig* & Ad Aertsen

Department of Neurobiology and Biophysics, Institute of Biology III, Albert-Ludwigs-University, Schänzlestraße 1, D-79104 Freiburg, Germany

The classical view of neural coding has emphasized the importance of information carried by the rate at which neurons discharge action potentials. More recent proposals that information may be carried by precise spike timing^{1–5} have been challenged by the assumption that these neurons operate in a noisy fashion—presumably reflecting fluctuations in synaptic input⁶—and, thus, incapable of transmitting signals with millisecond fidelity. Here we show that precisely synchronized action potentials can propagate within a model of cortical network activity that recapitulates many of the features of biological systems. An attractor, yielding a stable spiking precision in the (sub)millisecond range, governs the dynamics of synchronization. Our results indicate that a combinatorial neural code, based on rapid associations of groups of neurons co-ordinating their activity at the single spike level, is possible within a cortical-like network.

Evidence is accumulating that cortical neurons *in vivo* are capable of producing action potentials with high temporal accuracy. In recordings of multiple single-neuron activity in behaving monkeys, precisely timed action potentials have been systematically related to stimuli and behavioural events, indicating that these instances of precise spike timing play a functional role^{1–3}. Independent evidence for precise spike timing in cortical neurons came from intracellular recordings *in vitro*^{4,5}. But can an instance of synchronous spiking, once it has occurred, be successfully propagated by subsequent groups of cortical neurons? Under which input conditions can a group of cortical neurons engage in precisely coordinated spike timing, and are such conditions feasible in the cortical network? How can we clarify and quantify the notions of 'well timed' and 'reliable', which gained such a prominent role in the on-going debate on temporal coding in the brain?^{7–10} Clearly, these questions must be resolved to determine whether cortical computation on the basis of precise spike timing is possible. Preliminary results have been presented in abstract form^{11,12}.

To address these questions, we have studied the fine-grained temporal response properties of the 'integrate-and-fire' neuron, a widely used class of model neurons capturing essential properties of

* Present addresses: Department of Nonlinear Dynamics, Max-Planck-Institut für Stöpfungsforschung, Göttingen, Germany (M.D.); Future Technology Research, Honda R&D Europe, Offenbach, Germany (M.-O.G.)

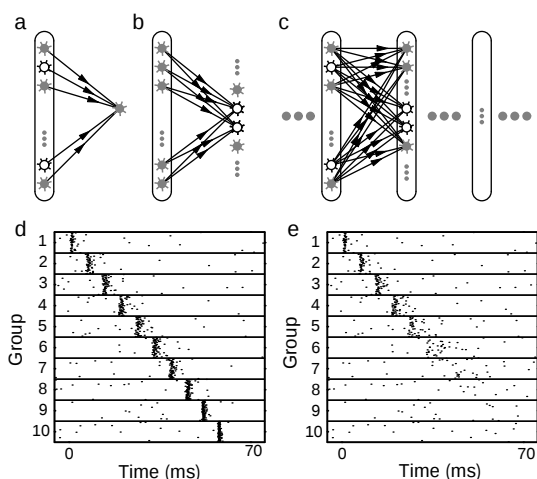


Figure 1 Propagation of synchronous spiking in cortical networks. Effect (a) and generation (b) of tight spike correlation by convergence of synchronous shared input onto target neurons (simultaneously recorded neurons are highlighted by black circles). c, Combination of a and b (no intragroup connections) generates synchronous spike volley. Propagation across neuron groups by repeating the arrangement ('synfire chain'¹⁷) can be stable (d) or unstable (e). d, e, Raster displays of propagating spike volley along fully connected synfire chain. Panels show the spikes in 10 successive groups of 100 neurons each (synaptic delays arbitrarily set to 5 ms). Initial spike volley (not shown) was fully synchronized ($\sigma = 0$ ms), containing $a = 50$ (d) or $a = 48$ (e) spikes; background activity identical in both cases.

the firing behaviour of cortical neurons^{13,14}. We focused on spike responses to transient membrane-potential excursions, implied by the physiological findings¹⁻⁵. As a rule, such transients are explained by convergent inputs from simultaneously spiking neurons onto a target neuron (Fig. 1a). These transients, in turn, result in well timed response spikes in target neurons. Neurons that share a large enough pool of simultaneously discharging input cells tend to align their action potentials¹⁵ (Fig. 1b). By repeating this arrangement, a group of neurons can reproduce its synchronous input activity and act as the source of synchronous shared input to the following group (Fig. 1c). This idea of connecting groups of neurons into feed-forward arrangements¹⁶ was formalized by Abeles in the 'synfire chain'¹⁷. The degree of temporal accuracy of spike times among the group's members determines whether subsequent groups can reproduce (or even improve) this accuracy (Fig. 1d), or whether synchronous excitation disperses and eventually dies out (Fig. 1e). Thus, in the context of cortical network function, the timing precision of a neuron's action potentials is measured *vis-à-vis* the timing of those of its companion neurons; the quality of timing is judged on the effect the group's activity has in the network, that is, whether synchronous spiking is sustained or whether it dies out.

Existing measures of short-term dynamics in neural transmission focus on two extreme cases of input activity: full synchrony and random arrival of spikes¹⁸⁻²⁰. Intermediate cases with limited degree of temporal dispersion are generally not addressed. 'Pulse packets'^{11,12} were introduced to overcome this restriction and to quantify the degree of synchrony in a propagating spike volley. A pulse packet characterizes a spike volley by two parameters: activity, a , and temporal dispersion, σ . Activity is defined as the number of spikes in the volley; their temporal dispersion is measured by the standard deviation of the underlying pulse density (Fig. 2a). Thus, in simulations we measured the response of a cortical model neuron (see Methods) to different pulse-packet inputs in the presence of background activity (Fig. 2b). The neuronal transmission function for transient input activity is defined by the transformation of the input pair (a_{in}, σ_{in}) into the output pair (α, σ_{out}), where α is the single neuron response probability. The firing probability curves (Fig. 2c)

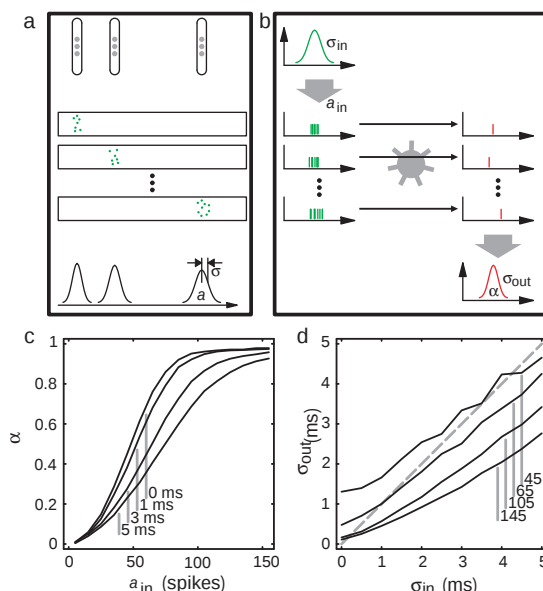


Figure 2 Neural transmission function for pulse-packet input. a, Pulse-packet representation of propagating spike volley. b, Cortical model neuron, stimulated with pulse packets by repeatedly drawing a_{in} spike times from Gaussian pulse density with standard deviation σ_{in} (green). Pulse packets were embedded in random spikes, mimicking background activity. The neuron responded with one spike at most, its timing varying trial by trial. The temporal distribution was measured by trial averaging (red). Net area under curve, spike probability α ; standard deviation, response jitter σ_{out} . c, Spike probability versus input spike number for different degrees of input synchrony. d, Temporal accuracy of spike response versus input synchrony for different input spike numbers. The diagonal (dashed grey line) marks points with equal temporal precision of input and response.

resemble the well known sigmoid activation function, the slope is determined by the degree of input synchrony (the saturation slightly below 1 reflects the chance that the neuron was refractory on stimulus arrival). As expected, the spread of the response distribution (Fig. 2d) increases with the input spread; however, the slope is less than 1, hence output spread increases more slowly than input spread. In addition, the curves show an offset; even for fully synchronized volleys ($\sigma_{in} = 0$ ms), some residual jitter of the response spike remains, reflecting the influence of background activity. Hence, each dispersion curve crosses the diagonal at some critical value of input synchrony. Up to this intersection, the neuron's response is less precise than the input, that is, synchronous input is desynchronized. Beyond the intersection, however, the neuron's spike response is more precise than the input, that is, the neuron exhibits a synchronizing behaviour. Duplication of the experiment with identical input pulse-packet realizations across trials yielded essentially the same results, confirming that trial-by-trial response variability is due to fluctuations in background activity²¹.

We used this neuronal transmission function to test whether the cortical network is capable of sustaining synchronous spiking activity. As each neuron responds to an incoming pulse packet with at most one spike, stable propagation of synchronous spike volleys inevitably requires the activation of successive, large enough groups of neurons (Fig. 1c). For a group of identical independent neurons, the distribution of response spikes to an input pulse packet is identical to the response distribution for a single neuron (Fig. 2). Thus, the spread of the group's response equals the single neuron's response dispersion σ_{out} (Fig. 2d). The expected number of response spikes a_{out} in a group equals α (Fig. 2c) multiplied by the group size w . Figure 3a, b (grey curves) shows the input-output relation for a group of $w = 100$ neurons. Assuming that the group's response to a

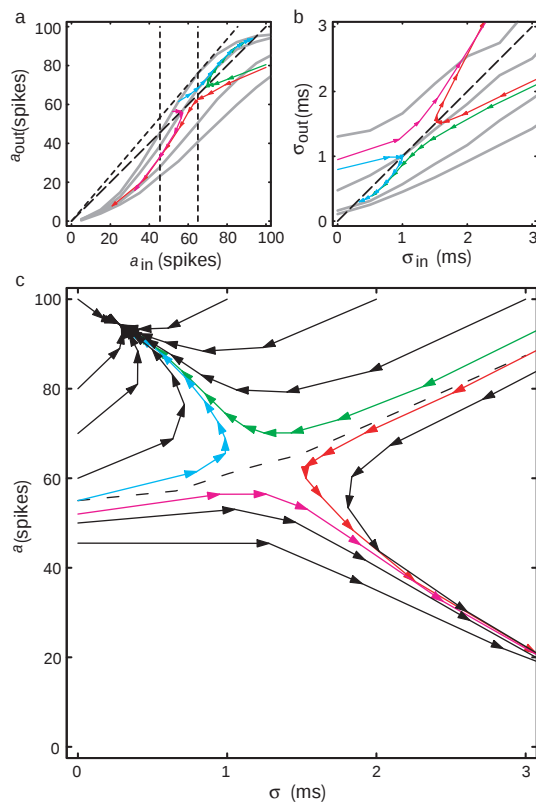


Figure 3 State-space analysis of propagating spike synchrony. **a, b**, Grey curves, transmission function for pulse-packet input for a group of $w = 100$ neurons. Coloured curves, evolution of synchronous spike volley along a chain of such groups (Fig. 1c) for different input configurations. Arrows, group to group transformation. Two cases of stable propagation are shown, one starting with few, fully synchronized spikes (blue; see Fig. 1d), the other with many, weakly synchronized spikes (green), and two unstable cases where synchronous activity rapidly dies out (red, purple; see Fig. 1e). **c**, State space portrait of dynamic variables a and σ . Coloured trajectories portray the four cases in **a, b**. The stable (upper) and unstable (lower) regimes are separated by the separatrix (dashed line).

spike volley is adequately described by a and σ (see Methods: Fluctuations), we can capture the evolution of synchronous spiking activity by repeatedly applying the (a, σ) -transformation while moving from one group to the next (Fig. 3a, b, coloured curves). Thus, the evolution of synchronous activity on its course through the network is described by a trajectory in the 2-dimensional state space, spanned by a and σ (Fig. 3c).

The state-space portrait exhibits two fixpoints: an attractor ($a = 95$, $\sigma = 0.5$ ms), and a saddle point ($a = 65$, $\sigma = 1.2$ ms). A separatrix (dashed line) running through the saddle point divides

the state space into two regimes. In the basin of attraction, all trajectories converge into the attractor. A spike volley starting anywhere inside this regime rapidly (that is, after only few stages) reaches a stable configuration (~ 95 spikes) with submillisecond dispersion. Volleys starting outside the stable regime decay after only few stages; too weak or too dispersed activity rapidly dies out. Note that neither the relationship between input and output activity (Fig. 3a) nor that between input and output jitter¹⁰ (Fig. 3b) alone determines whether synchronous activity survives. The nonmonotonic evolution of these variables along a trajectory (Fig. 3c, coloured curves) demonstrates this fact. An initial increase in temporal spread may still support stable propagation, provided that the number of spikes in the volley is large enough (blue). If, however, this number is too small, the volley dies out, in spite of its initial increase (purple). Conversely, synchronous activity may still vanish with an initial decrease in dispersion (red), unless the volley is large enough (green). Thus, the system dynamics are governed by the interaction of the two state variables.

Evidently, the number of neurons per group influences the evolution of the activity (see Fig. 3a). The analysis in Fig. 3 was made using the group size $w = 100$. To determine how many simultaneously firing neurons are needed to guarantee that synchronous activity survives in the network, we examined how the structure of the state space depends on the groups size (Fig. 4). For increasing numbers of neurons per group, the two fixpoints move apart, thereby increasing the basin of attraction (yellow). The regime over which synchronous spiking survives in the network increases accordingly. By contrast, for decreasing groups size, the two fixpoints approach each other until, at some critical value (here $w = 89$), they merge into a single saddle node. Below this critical value, no fixpoint exists and, hence, all trajectories lead to extinction (Fig. 4a). Thus, a minimum of some 90 neurons per group is needed to maintain precise spike synchrony. This lower bound is essentially determined by the ratio of the distance from mean membrane potential to spike threshold and the postsynaptic potential (PSP) amplitude; stronger intergroup synapses reduce this number (see Methods: Fluctuations).

Our results show that for a wide range of amplitudes and dispersions of spike input distributions, the response of successively activated groups of cortical neurons is governed by an attractor, which describes a stationary configuration of activity in (a, σ) space. However, unlike the Hopfield attractor²², this attractor describes a dynamic activity configuration in neuron space, that is, different neuron groups, one after the other, contribute single spikes to the propagating synchronous wave. The basin of attraction guarantees robustness of the propagating synchrony against perturbations exceeding the response variability accounted for by the transmission function (Fig. 2). Thus, temporal dispersion due to differences in axonal or dendritic delays, fluctuations in synaptic transfer properties or correlated background fluctuations will not destroy the synchronous transmission, as long as they do not push the network outside the basin of attraction.

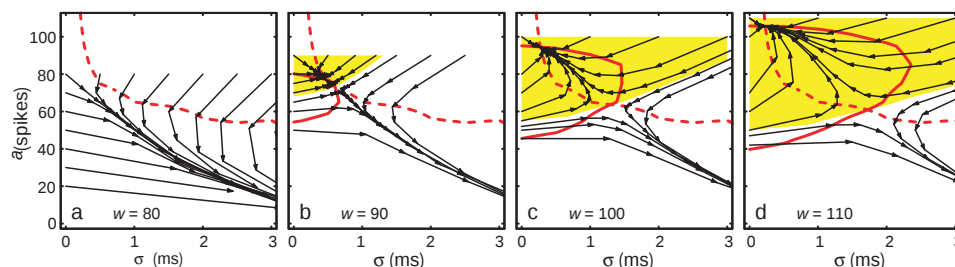


Figure 4 Dependence of the spike synchrony state space on the size of the neuron groups. **a–d**, State-space portraits of propagating spike synchrony for four different group sizes w , increasing from $w = 80$ (**a**) in steps of 10 up to $w = 110$ (**d**). In each

panel, the continuous red line indicates the a -isocline, the red dashed line indicates the σ -isocline (see Methods: Isoclines). The intersections of the two isoclines define the fixpoints. The basin of attraction is indicated in yellow.

In the stable state, essentially all response spikes in a volley fall within ± 1 ms. This temporal precision is consistent with the accuracy of observed spike patterns in cortical recordings^{1–3}. Thus, in contrast to other studies using this neuron model^{7,8}, we conclude that precise synchronous firing of cortical neurons is indeed feasible, in spite of the membrane time constant of 10 ms or more (see Methods: Membrane time constant). The notion of pulse packets yields a natural solution to the question of whether the cortical neuron acts as ‘integrator’ or ‘coincidence detector’—a question raised many years ago¹⁸ and revived recently^{7–10}—by embedding these two conceptions in a single framework. The temporal structure of the input determines which of the two aspects is emphasized.

For each neuron in our network model, the synapses delivering activity from the preceding group and the synapses delivering background activity were set at equal strength. Our findings show that stable transmission of synchronous spiking does not require dedicated strong synapses, provided enough neurons can be recruited in successive groups. Each neuron in a group contributes a single spike to the passing volley. Once the neuron recovers from refractoriness, it is ready to engage in another group. Thus, each neuron may participate—spike by spike, not limited by a specific arrangement of synaptic weights—in multiple volleys with different neuron compositions, provided its engagements differ by more than the refractory period. Hence, the network only needs to be locally feed forward. Several such volleys may propagate through the network simultaneously, allowing multiple synchronous processes to coexist while maintaining their identities^{23,24}. Their degree of temporal coupling may differ, depending on overlap (numbers and arrangement of shared neurons and inputs)^{23,25}. This scheme indicates that a combinatorial neural code, based on the continual reconfiguration of the cortical network into short-lived functional groups depending on the immediate computational demands, is compatible with biological constraints. □

Methods

Model neuron

Simulations were performed using a leaky-integrator with voltage-threshold model^{13,14}, with physiological and anatomical parameters taken from experimental literature. The model neuron (membrane time constant 10 ms, resting potential -70 mV, spike threshold -55 mV, absolute refractoriness 1 ms, relative refractoriness ($\tau \approx 15$ ms) modelled by K-conductances) was supplied with synaptic noise input, reflecting on-going activity in the cortical network (20,000 synapses: 88% excitatory, 12% inhibitory)²⁶. Postsynaptic currents (PSCs) were modelled by an α -function to yield realistic PSPs (peak amplitude 0.14 mV, time-to-peak 1.7 ms, half-width 8.5 ms)²⁷. Identical values were used for inter-group and background connections; excitatory and inhibitory PSPs only differed in sign. Background firing rates (excitatory, 2 Hz; inhibitory, 12.5 Hz; all uncorrelated stationary Poisson) were chosen to yield an output rate of 2 Hz. At this consistency condition, output statistics were approximately Poisson, membrane potential shot noise (mean 8.25 mV, s.d. 2.85 mV) was close to ‘balanced’ excitation/inhibition²⁸. It can be shown that details of the construction of background fluctuations are not essential. Simulations were performed in 0.1 ms time steps using the simulation tool SYNOD²⁹.

Threshold packet

The (a_{in}, a_{out}) curves (grey) in Fig. 3a cross the diagonal (dashed) in two points. Here, the number of output spikes equals the number of input spikes. The lowest intersection (at the left-most vertical line) occurs with the curve for fully synchronized input ($\sigma_{in} = 0$), defining a lower boundary on the size of the threshold packet, that is, the minimum input spike number needed to reach the attractor. Smaller packets cannot survive, since for other curves, the output spike number is even smaller. Decreasing the group’s size rotates the diagonal relative to the (a_{in}, a_{out}) curves counterclockwise around the origin; the intersection points approach each other and the threshold packet size increases. Thus, for decreasing w , more curves fall below the diagonal until even the curve for fully synchronized input only touches (short-dashed oblique line intersection with right-most vertical line). For smaller groups, all curves run below the diagonal; the attractor vanished, and stable propagation of synchronous spiking is no longer possible.

Isoclines

The a -isocline is the collection of states for which the spike number in a volley does not change from stage to stage, irrespective of σ (Fig. 4b–d, solid red curves). The σ -isocline contains all states maintaining temporal spread, irrespective of a (Fig. 4a–d, dashed red curves). Thus, the isoclines are the loci of horizontal/vertical flow. The fixpoints (neither a

nor σ changes) are the intersections of the isoclines. The σ -isocline is independent of w (Fig. 3b). The spike number, however, is proportional to w (Fig. 3a). Hence, with decreasing w (Fig. 4d to 4a), the a -isocline shrinks and moves left until it ceases to exist. Thus, the group’s size w acts as bifurcation parameter, controlling the existence and separation of fixpoints.

Fluctuations

As confirmed by network simulations (M.-O.G., manuscript in preparation), the state-space portrait derived here describes the evolution of synchronous activity in the mean, that is, by subsequent values of the expectation $\langle a, \sigma \rangle$ across trials with different background activity realizations. Around each point of a trajectory, these realizations form a distribution with width determined by a, σ, w and intergroup connectivity. This width becomes more important near the separatrix because of the increased probability—even for trajectories stable in the mean—that individual realizations leave the basin of attraction (and vice versa). Upscaling the synaptic weights by a factor up to 10 while downsampling the groups size accordingly does not alter the structure of the state space, except close to the separatrix where the probability to leave the basin of attraction increases. The contribution of individual input spikes grows; consequently, fluctuations in membrane-potential response to pulse-packet realizations with identical parameters and, hence, trial-by-trial variability, increases.

Background activity

Background activity in different neurons was considered independent, stationary Poisson. However, on-going cortical activity is known to exhibit coherent spatio-temporal structure³⁰. Hence, anatomically nearby neurons within a group tend to be excited (inhibited) together. This affects the pulse-packet properties needed to make these neurons fire simultaneously. The impact of such coherence in background activity is currently being studied.

Membrane time constant

The temporal precision of spike response is not constrained by membrane time constant itself; the limiting factor is the up slope of the excitatory PSP. The larger this slope, the faster the membrane-potential response to a pulse packet traverses the threshold region¹⁷. This reduces the chance of interference with background fluctuations, which degrade response spike precision. The membrane time constant does, however, limit transmission of synchronous spikes in the opposite way. It determines the integration time window of the receiving neuron, limiting the extent over which a spike volley is ‘seen’ as a single packet, rather than as individual spikes. For a too-small membrane time constant, PSPs no longer overlap and cannot add up to threshold. Reliable transmission of incompletely synchronized spike volleys, therefore, requires a minimal (rather than maximal) membrane time constant, in the order of the volley duration.

Received 17 June; accepted 23 August 1999.

1. Abeles, M., Bergman, H., Margalit, E. & Vaadia, E. Spatiotemporal firing patterns in the frontal cortex of behaving monkeys. *J. Neurophysiol.* **70**, 1629–1638 (1993).
2. Prut, Y. *et al.* Spatiotemporal structure of cortical activity: Properties and behavioral relevance. *J. Neurophysiol.* **79**, 2857–2874 (1998).
3. Riehle, A., Grün, S., Diesmann, M. & Aertsen, A. Spike synchronization and rate modulation differentially involved in motor cortical function. *Science* **278**, 1950–1953 (1997).
4. Mainen, Z. F. & Sejnowski, T. J. Reliability of spike timing in neocortical neurons. *Science* **268**, 1503–1506 (1995).
5. Nowak, L. G., Sanchez-Vives, M. V. & McCormick, D. A. Influence of low and high frequency inputs on spike timing in visual cortical neurons. *Cereb. Cortex* **7**, 487–501 (1997).
6. Calvin, W. H. & Stevens, C. F. Synaptic noise and other sources of randomness in motoneuron interspike intervals. *J. Neurophysiol.* **31**, 574–587 (1968).
7. Shadlen, M. N. & Newsome, W. T. Noise, neural codes and cortical organization. *Curr. Opin. Neurobiol.* **4**, 569–579 (1994).
8. Shadlen, M. N. & Newsome, W. T. The variable discharge of cortical neurons: Implications for connectivity, computation, and information coding. *J. Neurosci.* **18**, 3870–3896 (1998).
9. König, P., Engel, A. K. & Singer, W. Integrator or coincidence detector? The role of the cortical neuron revisited. *Trends Neurosci.* **19**, 130–137 (1996).
10. Marsalek, P., Koch, C. & Maunsell, J. On the relationship between synaptic input and spike output jitter in individual neurons. *Proc. Natl Acad. Sci. USA* **94**, 735–740 (1997).
11. Diesmann, M., Gewaltig, M.-O. & Aertsen, A. in *Computational Neuroscience—Trends in Research 1995* (ed. Bower, J.) 59–64 (Academic, San Diego, 1996).
12. Diesmann, M., Gewaltig, M.-O. & Aertsen, A. in *From Membrane to Mind* (eds Elsner, N. & Wässle, H.) 62 (Thieme, Stuttgart, 1997).
13. Gerstein, G. L. & Mandelblat, B. Random walk models for the spike activity of a single neuron. *Biophys. J.* **4**, 41–68 (1964).
14. Tuckwell, H. C. *Introduction to Theoretical Neurobiology* (Cambridge Univ. Press, Cambridge, 1988).
15. Perkel, D. H., Gerstein, G. L. & Moore, G. P. Neuronal spike trains and stochastic point processes. II. Simultaneous spike trains. *Biophys. J.* **7**, 419–440 (1967).
16. Griffith, J. S. On the stability of brain-like structures. *Biophys. J.* **3**, 299–308 (1963).
17. Abeles, M. *Corticonics* (Cambridge Univ. Press, Cambridge, 1991).
18. Abeles, M. Role of cortical neuron: integrator or coincidence detector? *Isr. J. Med. Sci.* **18**, 83–92 (1982).
19. Bernander, O., Koch, C. & Usher, M. The effect of synchronized inputs at the single neuron level. *Neural Comp.* **6**, 622–641 (1994).
20. Murthy, V. N. & Fetz, E. E. Effects of input synchrony on the firing rate of a three-conductance cortical neuron model. *Neural Comp.* **6**, 1111–1126 (1994).
21. Arieli, A., Sterkin, A., Grinvald, A. & Aertsen, A. Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. *Science* **273**, 1868–1871 (1996).

22. Hopfield, J. J. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl Acad. Sci. USA* **79**, 2554–2558 (1982).
23. Bienenstock, E. A model of neocortex. *Network* **6**, 179–224 (1995).
24. Herrmann, M., Hertz, J. A. & Prügel-Bennett, A. Analysis of synfire chains. *Network* **6**, 403–414 (1995).
25. Arnoldi, H.-M. R. & Brauer, W. Synchronization without oscillatory neurons. *Biol. Cybern.* **74**, 209–223 (1996).
26. Braitenberg, V. & Schüz, A. *Anatomy of the Cortex* (Springer, Berlin, 1991).
27. Fetz, E., Toyama, K. & Smith, W. in *Cerebral Cortex* Vol. 9 (eds Peters, A. & Jones, E. G.) 1–47 (Plenum, New York, 1991).
28. van Vreeswijk, C. & Sompolinsky, H. Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science* **274**, 1724–1726 (1996).
29. Diesmann, M., Gewaltig, M.-O. & Aertsen, A. *SYNOD: An Environment for Neural Systems Simulations* Technical report GC-AA/95-3 (The Weizmann Institute of Science, Rehovot, Israel, 1995). (<http://www.synod.uni-freiburg.de>).
30. Arieli, A., Shoham, D., Hildesheim, R. & Grinvald, A. Coherent spatiotemporal patterns of ongoing activity revealed by real-time optical imaging coupled with single-unit recording in the cat visual cortex. *J. Neurophysiol.* **73**, 2072–2093 (1995).

Acknowledgements

We thank M. Abeles, E. Bienenstock, S. Grün, I. Nelken, A. Riehle, S. Rotter and C. von der Malsburg for their constructive comments. Supported in part by grants from the Deutsche Forschungsgemeinschaft, the German-Israeli Foundation for Scientific Research and Development, and Human Frontier Science Program.

Correspondence and requests for materials should be addressed to A.A.
(e-mail: aertsen@biologie.uni-freiburg.de).

Membrane-anchored aspartyl protease with Alzheimer's disease β -secretase activity

Riqiang Yan^{*†}, Michael J. Bienkowski^{†‡}, Mary E. Shuck[‡], Huiyi Miao^{*}, Monica C. Tory[‡], Adele M. Pauley[§], John R. Brashler^{||}, Nancy C. Stratman[§], W. Rodney Mathews[¶], Allen E. Buhl^{||}, Donald B. Carter[§], Alfredo G. Tomasselli[#], Luis A. Parodi[‡], Robert L. Heinrikson[#] & Mark E. Gurney[§]

^{*} Cell & Molecular Biology, [‡] Genomics, [#] Protein Sciences, ^{||} Pharmacology, [¶] Structural, Analytical & Medicinal Chemistry and [§] Neurobiology, Pharmacia & Upjohn, Inc., 301 Henrietta Street, Kalamazoo, MI 49007, USA

[†] Bioinformatics, Pharmacia & Upjohn, Inc., Lindhagensgatan 133, S-11287 Stockholm, Sweden

[†] These authors contributed equally to this work

Mutations in the gene encoding the amyloid protein precursor (APP) cause autosomal dominant Alzheimer's disease^{1–3}. Cleavage of APP by unidentified proteases, referred to as β - and γ -secretases^{4–7}, generates the amyloid β -peptide, the main component of the amyloid plaques found in Alzheimer's disease patients⁸. The disease-causing mutations flank the protease cleavage sites in APP and facilitate its cleavage. Here we identify a new membrane-bound aspartyl protease (Asp2) with β -secretase activity. The Asp2 gene is expressed widely in brain and other tissues. Decreasing the expression of Asp2 in cells reduces amyloid β -peptide production and blocks the accumulation of the carboxy-terminal APP fragment that is created by β -secretase cleavage. Solubilized Asp2 protein cleaves a synthetic APP peptide substrate at the β -secretase site, and the rate of cleavage is increased tenfold by a mutation associated with early-onset Alzheimer's disease in Sweden³. Thus, Asp2 is a new protein target for drugs that are designed to block the production of amyloid β -peptide and the consequent formation of amyloid plaque in Alzheimer's disease.

Visual inspection suggests that the β - and γ -secretase cleavage sites in APP might be substrates for cleavage by aspartyl proteases,

and indeed, cathepsin D cleaves synthetic β -secretase substrates⁹. This cleavage is facilitated by the KM $\rightarrow$ NL mutation, referred to as the 'Swedish' mutation, found in patients with early-onset Alzheimer's disease¹⁰; however, APP processing to amyloid β (A β) peptides occurs normally in hippocampal neurons cultured from cathepsin-D-null mice¹¹. Nevertheless, it seemed plausible that the APP β - or γ -secretases could be as yet uncharacterized aspartyl proteases; therefore, we searched for new human enzymes of this mechanistic set. Sequencing of the *Caenorhabditis elegans* genome was nearing completion, which offered the possibility of enumerating the complete set of aspartyl proteases encoded in a simple metazoan genome, and using these as a bridge to human sequence databases.

Simple AWK scripts scanning for the D(S/T)G active-site motif, PROSITE and hidden Markov models were used to search the WormPep database of predicted *C. elegans* proteins. This revealed at least 10 candidate aspartyl proteases. Seven of these ten were found on a single chromosome, chromosome V (F21F8.3, F21F8.4, F21F8.7, Y39B6B.G, Y39B6B.J, Y39B6B.H and T18H9.2), and three each of these were found in the same cosmid clones (F21F8 and Y39B6B), suggesting that they represent a recently evolved family of proteins that arose by ancestral gene duplication. Other homologous predicted genes were found in the same cluster (F21F8.2, F21F8.6 and Y39B6B.I); however, these contain only a single DTG or DSG motif. Additional predicted aspartyl protease genes were found on chromosomes IV (C11D2.2) and X (R12H7.2 and H22K11.1). Searches of vertebrate expressed sequence tag (EST) databases with the 10 *C. elegans* sequences identified 7 known and 4 new candidate aspartyl proteases. The new human sequences were numbered in order of their discovery (Asp1–4). R12H7.2 and H22K11.1 appear to be *C. elegans* homologues of cathepsin D. Most of the chromosome V aspartyl proteases had no clear vertebrate orthologues; however, one of these (T18H9.2) bridged to two unusual sequences (Asp1 and Asp2) which contained the less common DSG motif in the second active site. In turn, C11D2.2 identified two additional sequences (Asp3 and Asp4) which have since been reported in the literature as napsins A and B¹².

The two predicted aspartyl protease sequences identified by T18H9.2 were of greatest interest. Completion of their sequences by a combination of EST sequencing, 5' rapid amplification of complementary DNA ends by the polymerase chain reaction, and library screening showed that both Asp1 and Asp2 had an unusual C-terminal extension containing a single predicted transmembrane domain (Fig. 1). Asp1 maps to human chromosome 21q22 within the Down's syndrome critical region, and Asp2 to chromosome 11q23–24. Northern hybridization to human tissue blots showed widespread expression of both Asp1 and Asp2. Both are expressed at the highest levels in pancreas. Asp2 is also expressed at high levels in brain, whereas Asp1 is expressed in brain at somewhat lower levels. *In situ* hybridization showed expression of Asp2 primarily in acinar cells of the exocrine pancreas, whereas faint hybridization was seen over neurons in hippocampus; however, we identified two Asp2 EST in a human astrocyte cDNA library indicating that Asp2 may be expressed in both neurons and glial cells. Transcripts for both Asp1 and Asp2 were expressed in human embryonic kidney 293 cells, human IMR-32 neuroblastoma cells and mouse Neuro-2a neuroblastoma cells, three commonly used cellular models of APP processing.

We used a panel of antisense oligomers to test the involvement of each of the four predicted aspartyl proteases in APP processing by a stable clone of HEK293 cells that had been engineered to process APP to A β peptides at high levels. These cells were transformed with a modified human APP695 cDNA containing the Swedish KM $\rightarrow$ NL mutation to which two lysine residues had been added to the C terminus (HEK/APP-Sw-KK cells). The KK motif greatly increases the processing and release of A β peptides but does not influence the ratio of A β 42/(A β 42 + A β 40), nor alter the effect of

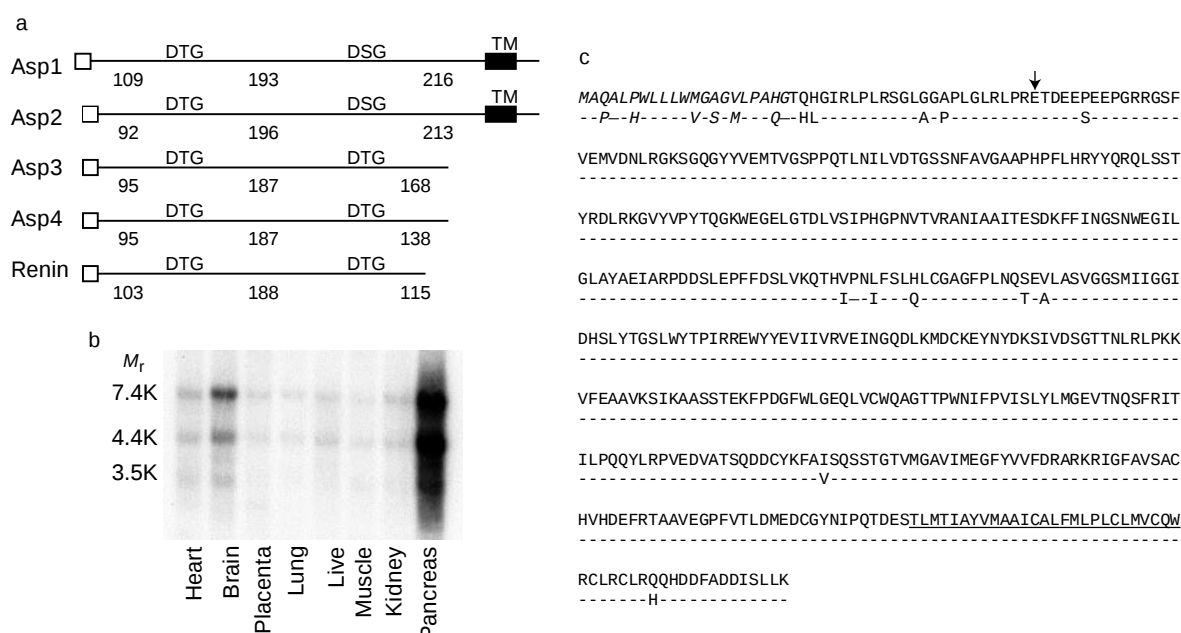


Figure 1 Asp2 functional domains, tissue distribution and amino-acid sequence. **a**, Alignment of the predicted primary structures of Asp1–Asp4 with human renin. Asp1 and Asp2 (51% amino-acid identity) have a DSG motif in the second predicted active site and possess an unusual C-terminal extension containing a single predicted transmembrane domain (TM). The predicted signal peptide for secretion (open box) and the number of residues in each domain is indicated. **b**, Tissue distribution of human Asp2.

the KM → NL or V717F mutations. These mutations increase total Aβ processing^{4,5} or selectively increase the production of Aβ42⁶, respectively. Thus, the HEK/APP-Sw-KK cells provide a sensitized background on which to screen for inhibition of Aβ processing.

Transfection of HEK/APP-Sw-KK cells with the panel of 16 antisense oligomers (four each targeting Asp1–Asp4) showed that only those oligomers targeting Asp2 considerably decreased the release of Aβ peptides into the medium. Two of the Asp2 antisense oligomers were chosen for resynthesis, as well as for synthesis of two additional oligomers containing the reverse sequence for use as controls. Their effects on transfected HEK/APP-Sw-KK cells are shown in Fig. 2. Both of the antisense oligomers targeting the human Asp2 transcript reduced Asp2 message levels, whereas the control reverse oligomers did not. Both also reduced the release of Aβ peptides into the culture medium. The inhibition of Aβ release ranged from 50% to 80% in many separate experiments and probably was dependent upon transfection efficiency. The antisense oligomers reduced the production of both Aβ40 and Aβ42 by roughly the same amount. The reduction of Aβ peptide production also was confirmed by immunoprecipitation and western blotting. This indicates that Asp2 may be involved directly or indirectly in the production of Aβ peptides and their release from HEK293 cells.

Because HEK293 cells derive from kidney, we extended the experiment to human IMR-32 neuroblastoma cells, which express all three APP isoforms¹³ and which release Aβ peptides into conditioned medium at measurable concentrations¹⁴, and obtained essentially identical results. The Asp2-1A and Asp2-2A antisense oligomers reduced Asp2 messenger RNA by 75% and 39%, respectively (quantitated using a TaqMan probe), whereas the reverse control oligomers had no effect. Correspondingly, release of Aβ40 and Aβ42 was reduced by 49 ± 2% and 42 ± 14% from cells treated with Asp2-1A, and by 43 ± 3 and 44 ± 18 with Asp2-2A ($P < 0.001$). Again, the reverse control oligomers had no effect.

Similar results were obtained in a murine system. Molecular cloning of the mouse Asp2 cDNA revealed a remarkable 98% amino-acid identity to human (Fig. 1c) and complete nucleotide

mRNA expression as shown by northern hybridization. The relative molecular mass (M_r) of each transcript is indicated. **c**, Clustal W sequence alignment of Asp2 from human (top line) and mouse (second line). The signal peptide is indicated in italics, the predicted transmembrane domain is underlined, and the active-site sequences are in bold. Arrow indicates the N terminus of purified recombinant Asp2 expressed in CHO cells.

identity at the sites targeted by the Asp2-1A and Asp2-2A antisense oligomers. In mouse Neuro-2A cells engineered to express APP-Sw-KK, the Asp2-1A antisense oligomer reduced the release of Aβ40 and Aβ42 by 70 ± 7% and 67 ± 2%, whereas a reduction of 61 ± 12% was seen for the release of both Aβ40 and Aβ42 from cells treated with Asp2-2A ($P < 0.001$). The reverse control oligomers had no effect. Thus, the three antisense experiments with HEK293, IMR-32 and Neuro-2a cells indicate that Asp2 is directly or indirectly involved in Aβ processing in both somatic and neural cell lines.

Treatment of HEK293/APP-Sw-KK cells with the Asp2 antisense oligomers had little effect on the release of total soluble APP (sAPP) from cells although they did appear to alter the ratio of sAPP isoforms released by either α- or β-secretase cleavage (Fig. 3). These cleavages generate species of soluble APP (sAPPα and sAPPβ, respectively), which contain different C termini which can be distinguished by the 6E10 monoclonal antibody that recognizes Aβ residues 1–16. As shown in Fig. 3b, no change in the release of total sAPP is shown on western blots developed with the 22C11 monoclonal antibody that reacts with an amino-terminal epitope of APP, whereas development with 6E10 shows an increase in sAPPα release from cells treated with either of the Asp2 antisense oligomers. An enzyme-linked immunosorbent assay (EIA) specific for sAPPα showed that release increased at least twofold, from 2.8 μg ml⁻¹ to 6.7 μg ml⁻¹ ($P < 0.005$), in cells transfected with the Asp2-2 antisense oligomer.

Cleavage of APP at either the β- or α-secretase sites also leaves C-terminal fragments containing the APP transmembrane domain and cytoplasmic tail. These contain 99 and 83 amino acids (CTF99 and CTF83), respectively (Fig. 3a). Indeed, in HEK293/APP-Sw-KK cells treated with Asp2 antisense oligomers, the amount of the CTF99 β-secretase product is reduced (Fig. 3c). Correspondingly, co-transfection of HEK293 cells or Neuro-2A cells with human Asp2 and APP-KK increases the production of CTF99 in comparison with cells co-transfected with APP-KK and empty vector DNA. Production of CTF99 is increased still further in cells co-transfected

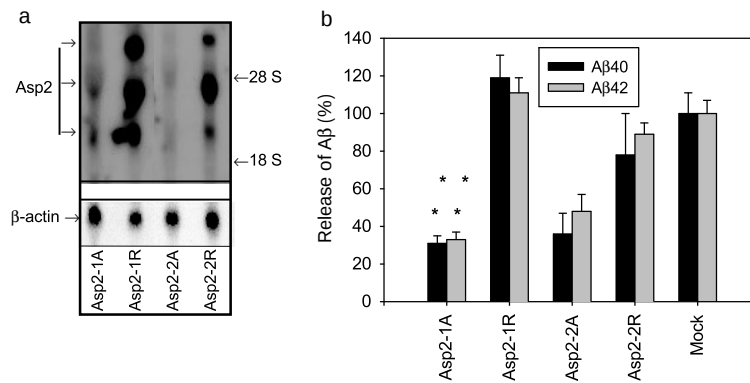


Figure 2 Asp2 antisense oligomers decrease amyloid β -peptide processing. **a**, Asp2 antisense oligomers (Asp2-1A and Asp2-2A) targeting two different sites on the Asp2 transcript specifically reduce Asp2 mRNA in transfected HEK/APP-Sw-KK cells as determined by northern hybridization, whereas control oligomers (Asp2-1R and Asp2-2R) with the reverse sequence lacked this effect. Arrows indicate the three Asp2 transcripts

with constructs expressing Asp2 and APP-Sw-KK. Thus, Asp2 appears specifically to facilitate β -secretase cleavage of APP and this effect is enhanced by the Swedish KM $\rightarrow$ NL mutation. This is in contrast to the effects of presenilin-1 gene disruption, which specifically increases the accumulation of CTF99 in cultured mouse neurons because of inhibition of γ -secretase cleavage^{15,16}.

Effects of Asp2 on the production of A β peptides from endogenously expressed APP isoforms were assessed in HEK293 cells transfected with a construct expressing Asp2 or with the empty vector after selection of transformants with the antibiotic G418. A β 40 production was increased in cells transformed with the Asp2 construct in comparison with those transformed with the empty vector DNA, the concentrations in conditioned medium were 424 ± 45 pg ml⁻¹ and 113 ± 58 pg ml⁻¹, respectively ($P < 0.001$). A β 42 release was below the limit of detection by the EIA, whereas the release of sAPP α was unaffected, 112 ± 8 ng ml⁻¹ versus 111 ± 40 ng ml⁻¹. These results provide further support for the hypothesis that Asp2 functions in the processing and release of A β from endogenously expressed APP.

present in HEK293 cells. No change was seen in β -actin mRNA in transfected cells.

b, Release of A β 40 and A β 42 from HEK/APP-Sw-KK cells, as measured by EIA, was reduced specifically by the Asp2 antisense oligomers (asterisks, $P < 0.001$). Antisense oligomers were transfected in quadruplicate cultures. Release was normalized against values for Mock-transfected cells treated with oligofectin-G only.

To determine the effects of Asp2 on the production of A β peptides from mutant APP, we transfected the two pools of cells with a panel of APP constructs. This showed that co-expression of Asp2 with APP or APP-VF increased A β 40 release from cells by 44% and 36%, respectively ($P < 0.05$), and that this effect was magnified by addition of the KK motif (126% and 186%, respectively, $P < 0.001$), Fig. 4a. Consistent with other reports⁶, the V717F mutation increased the production of A β 42 relative to total A β peptides, which was further increased to 175% by co-expression of Asp2 ($P < 0.001$), but there was no change in the relative ratio, Fig. 4b, c. Thus, Asp2 had little effect on the choice of γ -secretase cleavage sites in transfected HEK293 cells.

Co-expression of Asp2 in HEK293 cells with constructs containing the Swedish KM $\rightarrow$ NL mutation did not have the same effect on A β production. Consistent with other reports^{4,5}, the Swedish mutation increased the production of both A β 40 and A β 42 with no change in their relative ratio (Fig. 4). However, when Asp2 was overexpressed in HEK293 cells, co-expression of APP-Sw caused a decrease in A β 40 and A β 42 release in comparison with control

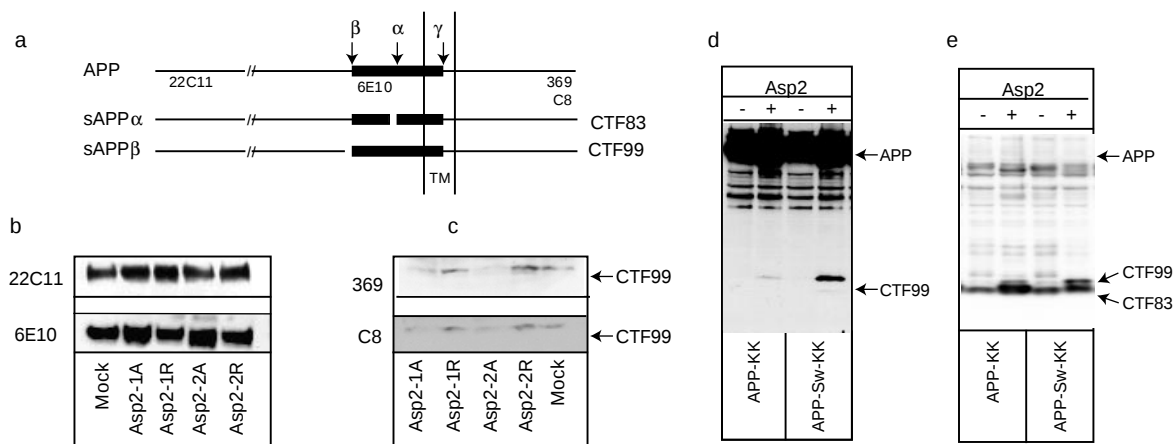


Figure 3 Asp2 increases production of the APP C-terminal β -secretase product. **a**, Illustration of the β -, α - and γ -secretase cleavage sites in APP and location of the 22C11, 6E10, 369 and C8 epitopes. The location of the A β peptide within APP is indicated (box). Processing at the α -secretase site cleaves the mid-region of the A β sequence and liberates the sAPP α ectodomain containing the 6E10 epitope, whereas the consequent 83-amino-acid C-terminal fragment (CTF83) retains the APP transmembrane domain (TM). Processing at the β -secretase site releases the sAPP β ectodomain and creates a 99-amino-acid C-terminal fragment (CTF99) containing the 6E10 epitope. **b**, Equal amounts of conditioned culture supernatants from HEK/APP-Sw-KK cells were analysed on western blots developed with the 22C11 and 6E10 antibodies. Cultures were

treated with the oligomers or transfection reagent as indicated. **c**, Equal amounts of protein from lysates of HEK/APP-Sw-KK cells were immunoprecipitated with either the 369 or C8 antibody as indicated, and analysed on western blots developed with 6E10 to identify CTF99. Cultures were treated with the oligomers or transfection reagent as indicated. **d**, Neuro-2A cells were co-transfected with either APP-KK or APP-Sw-KK, with or without Asp2 as indicated. Equal amounts of protein from cell lysates were analysed by western blot developed with 6E10 to detect APP and CTF99 (arrows). **e**, HEK293 were co-transfected with either APP-KK or APP-Sw-KK, with or without Asp2 as indicated. Equal amounts of protein from cell lysates were analysed by western blot developed with 369 antibody to detect APP, CTF99 and CTF83 (arrows).

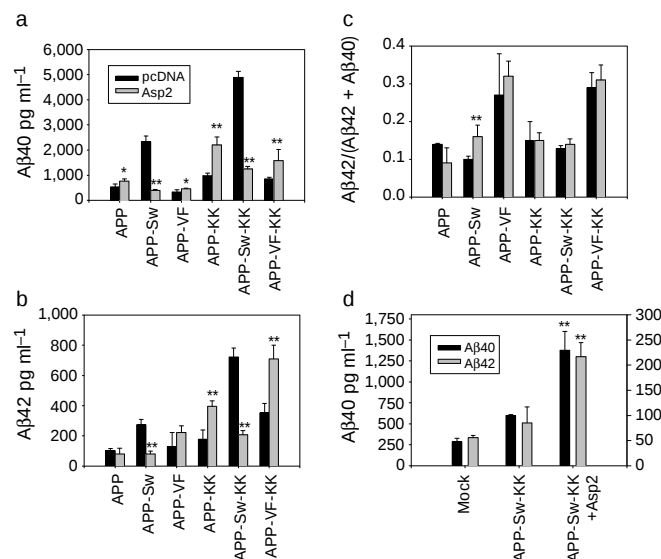


Figure 4 Asp2 increases A β peptide release. **a**, Production of A β 40 from HEK293 cells transfected with Asp2 or empty vector DNA (pcDNA) subsequently transfected with the indicated APP constructs. **b**, Production of A β 42 as in **a**. **c**, Ratio of A β 42/(A β 42+A β 40) produced as in **a**. The APP constructs were transfected in quadruplicate cultures (asterisk, $P < 0.05$; double asterisk, $P < 0.001$). **d**, Production of A β 40 and A β 42 by IMR-32 cells co-transfected with APP-Sw-KK and either Asp2 or empty vector DNA.

cells transformed with the empty vector. This effect is cell-line-dependent as co-transfection of IMR-32 cells with vectors expressing Asp2 and APP-Sw-KK results in more than a twofold increase in the production of both A β 40 and A β 42 in comparison with cells co-transfected with the APP-Sw-KK construct and empty vector DNA (Fig. 4d). Differences in the level of Asp2 expression or its localization within HEK293 or IMR-32 cells may account for this difference.

We obtained direct evidence that Asp2 possesses β -secretase activity using biochemical studies that measured purified Asp2 proteolytic activity against synthetic APP peptide substrates. Native, full-length Asp2 was expressed in Chinese hamster ovary (CHO) cells. Its behaviour on cell fractionation, detergent solubilization and purification by sequential chromatography was consistent with that of an integral membrane protein. Sequence analysis of the purified recombinant protein indicated a major N-terminal sequence beginning with a glutamic acid (arrow in Fig. 1c); however, at present, it is unclear whether this is the N terminus of mature Asp2.

Two peptides were designed for assaying β -secretase activity. The first contained the wild-type APP β -secretase site, whereas the second contained the Swedish KM $\rightarrow$ NL modification of the β -secretase cleavage site. Maximal activity was seen with the Swedish β -secretase peptide. As expected for an aspartyl protease, proteolytic activity was sensitive to pH with maximal hydrolysis seen at pH 5.0. Amino-terminal sequencing of the two cleavage products verified that cleavage occurred at the sequence NL $\downarrow$ DA (Fig. 5a). The rate of cleavage was reduced tenfold in the corresponding wild-type APP peptide (Fig. 5b). Proteolytic activity was insensitive to 8 μ M pepstatin or a mixture of 10 μ M leupeptin, 10 μ M E-64 and 5 mM EDTA, inhibitors of cathepsin D (and other aspartyl proteases), serine proteases, cysteinyl proteases, and metalloproteases, respectively. 'Mock' preparations of unmodified CHO cell membranes did not contain substantial Asp2-like activity. Thus, Asp2 acts directly in cell-free assays to cleave synthetic APP peptides at the β -secretase site, and the rate of cleavage is greatly increased by the Swedish KM $\rightarrow$ NL mutation associated with Alzheimer's disease.

Our experiments associated a new human aspartyl protease with the processing of APP at the β -secretase cleavage site. This protease

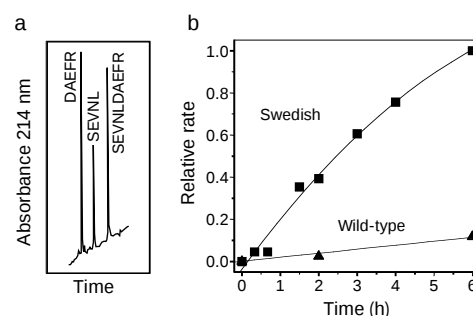


Figure 5 Asp2 β -secretase cleavage activity. **a**, Reverse phase HPLC profile showing the products of Asp2 cleavage. The amino-acid sequences of the parent peptide and the two hydrolysis products are indicated. **b**, Relative rates of hydrolysis of wild-type (triangles) and Swedish (squares) APP β -secretase peptides by Asp2.

contains an unusual C-terminal transmembrane domain that may help it colocalize with APP within cellular membranes where A β processing occurs. The expression pattern of Asp2 suggests a normal function in the brain, as well as in the exocrine pancreas. Our data indicate that Asp2 functions in the β -secretase pathway in cell lines of both somatic and neural origin, and that the enzyme meets many of the criteria expected of a APP β -secretase. Whether or not other enzymes also possess β -secretase activity is not excluded by these experiments. We note that identification of Asp2 as a candidate APP β -secretase has been reported independently by two other groups^{17,18}. Thus, inhibitors of Asp2 will provide a new approach to the treatment of Alzheimer's disease. □

Methods

Reagents

Northern hybridization was performed using human tissue blots (Clontech). Chromosomal localization was performed by Genome Systems, Inc. *In situ* hybridization to human tissues was performed by LifeSpan Biosciences, Inc. We used 6E10 and 4G8 (Senetek), 22C11 (Boehringer-Mannheim), LN27 (Zymed Laboratories) Rb162 and Rb165 (New York Institute for Basic Research), 369 (Paul Greengard), and C8 (Dennis Selkoe) antibodies. Oligofectin-G and the Asp2-1 (5'-CCCATAACAGTGCCTGGATGACT-3') and Asp2-2 (5'-GAACTCATCTGTCACATGGCAAGCG-3') chimeric antisense oligomers were from Sequitur, Inc. APP constructs were assembled in the vector pIRES (Clontech). Asp2 constructs were assembled in the vector pcDNA3.1/hygro (Invitrogen). The sequence for the HEX-labelled TaqMan probe (Perkin Elmer) was 5'-AGGGCAA CAACGACGCGCAATTACA-3'. Amplification was performed with the primer pair 5'-TCAGAGCAGCCAATGGCC-3' and 5'-GCCTGTAGGTGGCTGGACA-3'.

Transfection and immunodetection

Oligofectin-G was used for lipid-mediated transfection of antisense oligomers into cultured cells according to the manufacturer's protocol. Supernatants and cell lysates were harvested 72 h after transfection. Transfection of plasmid DNA was performed using either the calcium phosphate method or Lipofectamine (GIBCO-BRL). The total amount of DNA used for transfection was held constant by adding empty vector DNA to the transfection mixture. Cell lysates and supernatants were harvested 48–160 h after transfection. Cells were lysed with 10 mM HEPES, pH 7.9, 150 mM NaCl, 10% glycerol, 1 mM EGTA, 1 mM EDTA, 0.1 mM sodium vanadate and 1% NP-40. Equal amounts of protein (50 μ g) were resolved on 4–12% Tricine gels (Novex), and transferred to nitrocellulose membranes for probing with 6E10, 22C11 or C8 antibodies. For immunoprecipitation, equal amounts of protein corresponding to one plate of cells were incubated with C8 or 369 antibody at 4 °C overnight, captured with protein A/protein G agarose beads, and processed for western blot detection of CTF99 with 6E10. The A β EIA was done as described¹⁹ using 6E10 monoclonal antibody as a capture antibody and biotinylated Rb162 or Rb165 antibody for detection of A β 40 and A β 42, respectively. The sAPP α EIA used LN27 antibody as a capture antibody and biotinylated 6E10 for detection.

Proteolytic activity assays

Recombinant Asp2 was purified from CHO cell membranes by solubilization in 25 mM Tris-HCl, pH 8.0, containing 50 mM β -octylglucoside followed by sequential chromatography on MonoQ and MonoS columns. Material prepared in this manner was more than 95% pure by SDS-PAGE analysis. Activity assays for Asp2 were performed using synthetic peptide substrates containing either the wild-type APP β -secretase site (SEVKM $\downarrow$ DAEFR) or the Swedish KM $\rightarrow$ NL mutation (SEVNL $\downarrow$ DAEFR). Reactions were performed in 50 mM 2-(N-morpholino)ethane-sulfonate, pH 5.5, containing 70 μ M peptide substrate and recombinant Asp2 for various times at 37 °C. The reaction products were quantified by reverse phase HPLC. The sequence of both products was confirmed using Edman sequencing and MADLI-TOF mass spectrometry.

Received 19 October; accepted 5 November 1999.

- Goate, A. *et al.* Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. *Nature* **349**, 704–705 (1991).
- Murrell, J., Farlow, M., Ghetti, B. & Benson, M. D. A mutation in the amyloid precursor protein associated with hereditary Alzheimer's disease. *Science* **254**, 97–99 (1991).
- Mullan, M. *et al.* A pathogenic mutation for probably Alzheimer's disease in the APP gene at the N-terminus of beta amyloid. *Nature Genet.* **1**, 345–347 (1992).
- Cai, X. D., Golde, T. E. & Younkin, S. G. Release of excess amyloid beta protein from a mutant amyloid beta protein precursor. *Science* **259**, 514–516 (1993).
- Citron, M. *et al.* Mutation of the beta-amyloid precursor protein in familial Alzheimer's disease increases beta-protein production. *Nature* **360**, 372–374 (1992).
- Suzuki, N. *et al.* An increased percentage of long amyloid beta protein secreted by familial amyloid beta protein precursor (beta APP717) mutants. *Science* **264**, 1336–1340 (1994).
- Selkoe, D. J. The cell biology of beta-amyloid precursor protein and presenilin in Alzheimer's disease. *Trends Cell Biol.* **8**, 447–453 (1998).
- Glenner, G. G. & Wong, C. W. Alzheimer's disease: initial report of the purification and characterization of a novel cerebrovascular amyloidogenic derivative. *Science* **255**, 728–730 (1984).
- Lador, U. S., Snyder, S. W., Wang, G. T., Holzman, T. F. & Krafft, G. A. Cleavage at the amino and carboxyl termini of Alzheimer's amyloid-beta by cathepsin D. *J. Biol. Chem.* **269**, 18422–18428 (1994).
- Dreyer, R. N. *et al.* Processing of the pre-beta-amyloid protein by cathepsin D is enhanced by a familial Alzheimer's disease mutation. *Eur. J. Biochem.* **224**, 265–271 (1994).
- Saftig, P. *et al.* Amyloidogenic processing of human amyloid precursor protein in hippocampal neurons devoid of cathepsin D. *J. Biol. Chem.* **271**, 27241–27244 (1996).
- Tatnell, P. J. *et al.* Napsins: new human aspartic proteinases. Distinction between two closely related genes. *FEBS Lett.* **441**, 43–48 (1994).
- Neill, D., Hughes, D., Edwardson, B. K., Rima, B. K. & Allsop, D. Human IMR-32 neuroblastoma cells as a model cell line in Alzheimer's Disease research. *J. Neurosci. Res.* **39**, 482–493 (1994).
- Asami-Odaka, A., Ishibashi, Y., Kikuchi, T., Kitada, C. & Suzuki, N. Long amyloid β -protein secreted from wild-type human neuroblastoma IMR-32 cells. *Biochemistry* **34**, 10272–10278 (1995).
- De Strooper, B. *et al.* Deficiency of presenilin 1 inhibits cleavage of amyloid precursor protein. *Nature* **391**, 387–390 (1998).
- Naruse, S. *et al.* Effects of PS1 deficiency on membrane protein trafficking in neurons. *Neuron* **21**, 1213–1221 (1998).
- Vassar, R. *et al.* β -Secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE. *Science* **286**, 735–741 (1999).
- Hussain, I. *et al.* Identification of a novel aspartic protease (Asp2) as β -secretase. *Molec. Cell. Neurosci.* [online] (<http://www.apnet.com/www/journal/cn/mcne.1999.0811>)
- Pirttilä, T. *et al.* Longitudinal study of cerebrospinal fluid amyloid proteins and apolipoprotein E in patients with probably Alzheimer's disease. *Neurosci. Lett.* **249**, 21–24 (1998).

Acknowledgements

We thank C. Himes, M. Fairbanks, J. Leone, T. Emmons, R. Drong, J. Slighom, G. Winterowd and D. McKinley for their help, and J. McCall for his unflagging support and good humour.

Correspondence and requests for materials should be addressed to R.Y. (e-mail: riqiang.yan@am.pnu.com) or M.E.G. (e-mail: mark.e.gurney@am.pnu.com). Sequences are deposited in GenBank under the following accession numbers: human Asp1, AF200342; human Asp2, AF200343; mouse Asp2, AF200346; human Asp3, AF200343; and human Asp4, AF200345.

Purification and cloning of amyloid precursor protein β -secretase from human brain

Sukanto Sinha, John P. Anderson, Robin Barbour, Guriqbal S. Basi, Russell Caccavello, David Davis, Minhtam Doan, Harry F. Dovey, Normand Frigon, Jin Hong, Kirsten Jacobson-Croak, Nancy Jewett, Pamela Keim, Jeroen Knops, Ivan Lieberburg, Michael Power, Hua Tan, Gwen Tatsuno, Jay Tung, Dale Schenk, Peter Seubert, Susanna M. Suomensaaari, Shuwen Wang, Donald Walker, Jun Zhao, Lisa McConlogue & Varghese John

Elan Pharmaceuticals, 800 Gateway Boulevard, South San Francisco, California 94080, USA

Proteolytic processing of the amyloid precursor protein (APP) generates amyloid β ($A\beta$) peptide, which is thought to be causal for the pathology and subsequent cognitive decline in Alzheimer's disease. Cleavage by β -secretase at the amino terminus of the $A\beta$ peptide sequence, between residues 671 and 672 of APP,

leads to the generation and extracellular release of β -cleaved soluble APP¹, and a corresponding cell-associated carboxy-terminal fragment. Cleavage of the C-terminal fragment by γ -secretase(s) leads to the formation of $A\beta$. The pathogenic mutation K670M671 $\rightarrow$ N670L671 at the β -secretase cleavage site in APP², which was discovered in a Swedish family with familial Alzheimer's disease, leads to increased β -secretase cleavage of the mutant substrate³. Here we describe a membrane-bound enzyme activity that cleaves full-length APP at the β -secretase cleavage site, and find it to be the predominant β -cleavage activity in human brain. We have purified this enzyme activity to homogeneity from human brain using a new substrate analogue inhibitor of the enzyme activity, and show that the purified enzyme has all the properties predicted for β -secretase. Cloning and expression of the enzyme reveals that human brain β -secretase is a new membrane-bound aspartic proteinase.

β -cleaved soluble APP (β -sAPP) was detected in membranes isolated from 293 cells stably overexpressing the 'Swedish' mutation, SweAPP751, by western blot analyses using the β -cleaved soluble APP (β -sAPP)-specific antibody Sw192 (ref. 4). Incubation of the membranes at pH 5.5 led to an increase in the cell-associated β -sAPP, and the appearance of a faster migrating species (Fig. 1). Treatment with O-glycanase resulted in the co-migration of both immunoreactive bands at the size of the lower band, which indicated that the smaller species resulted from β -cleavage of membrane-associated N-glycosylated immature APP (data not shown). These results are consistent with the specific cleavage of the full-length APP at the β -cleavage site by a membrane-bound proteinase activity. The membrane-bound β -cleavage activity exhibited a preference for acidic pH, with an optimum value of pH 5.5. Co-incubation with class-specific protease inhibitors, such as pepstatin, E-64 or phenylmethylsulphonyl fluoride, did not affect the generation of the β -cleaved APP (data not shown). Washing the membranes with 0.1% saponin under hypotonic conditions did not lead to loss of the membrane-associated β -cleavage enzyme activity (Fig. 1); therefore, we extracted P2 membranes⁵ in 0.1% Triton X-100, 0.1% Brij-35 or 0.1% β -octylglucoside to test the solubility of enzyme activity. The soluble supernatant fractions were assayed for β -cleavage activity, on an exogenous recombinant substrate, MBPC125Swe. Specific β -cleavage was detected only in the Triton X-100 extracts.

We analysed various tissues and cell lines for β -cleavage activity, by extracting P2 membranes from each source with 0.2% Triton X-100 and assaying for β -cleavage (Fig. 2a). Human and mouse brain, and brain regions had uniformly high levels of enzyme activity, whereas little activity was detected in other tissues. In different cell lines, neurons had the highest level of enzyme activity, whereas 293, Cos and Chinese hamster ovary (CHO) cells had lower levels. Cells

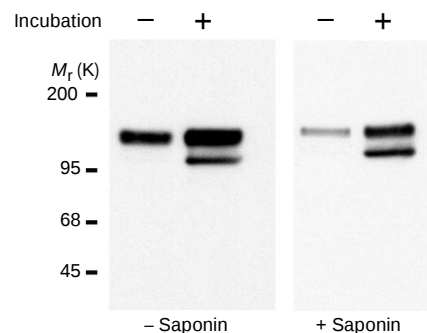


Figure 1 Endogenous substrate cleavage by β -secretase in P2 membranes. Membranes were prepared from 293 cells stably transfected with APP751, and either extracted with 0.1% saponin (+ saponin) or used directly (– saponin). Membranes in 0.1 M sodium acetate, pH 5.5 and 2% DMSO were either incubated (+) or solubilized without incubation (–). Samples were analysed by immunoblotting with the β -cleavage-specific 192sw antibody. M_r , relative molecular mass.

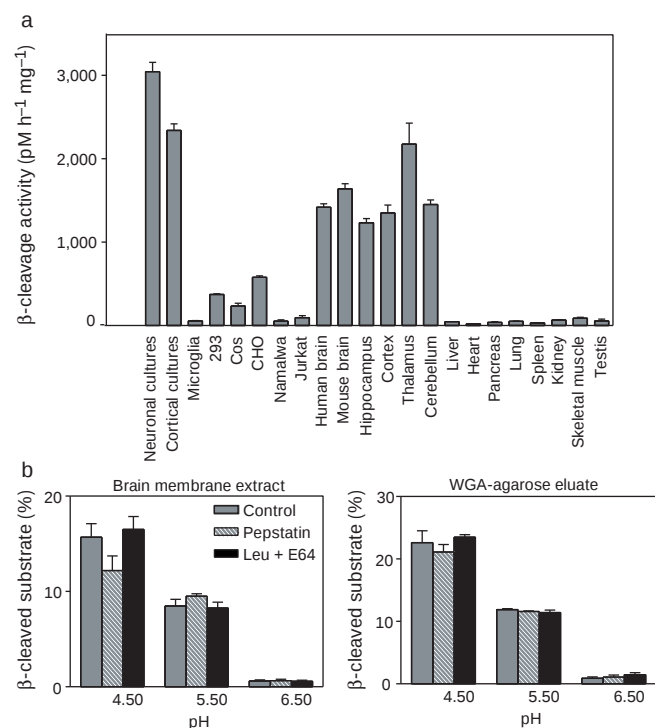


Figure 2 β -secretase activity in tissues and cell lines. **a**, Triton extracts of P2 membranes from the indicated cell types (first five samples) or tissues were assayed for β -secretase activity. **b**, Triton extracts of human brain membranes were assayed for β -secretase directly or after purification on WGA-agarose. Assays were performed at the indicated pH in the absence or presence of pepstatin or leupeptin and E-64 (Leu + E64).

of monocytic or lymphocytic origin did not have detectable enzyme activity. Thus, the enzyme activity is highest in cells of central nervous system (CNS) lineage, and found in cell lines commonly used for analysis of APP metabolism, in line with the observation that β -sAPP production is enhanced in CNS-derived cells, such as foetal neurons in culture¹.

An assay of solubilized extracts of human brain membranes showed a strong pH-dependent β -cleavage activity, which, like the activity in the 293Swe membranes, was not inhibited by pepstatin, leupeptin or E-64 (Fig. 2b). The soluble enzyme activity quantitatively bound to and eluted from wheat germ agglutinin (WGA) lectin-agarose. The pH dependence and inhibitor resistance remained unchanged through this (Fig. 2b) and all subsequent purification steps, and the enzyme activity migrated as a single species when analysed by size-exclusion chromatography (data not

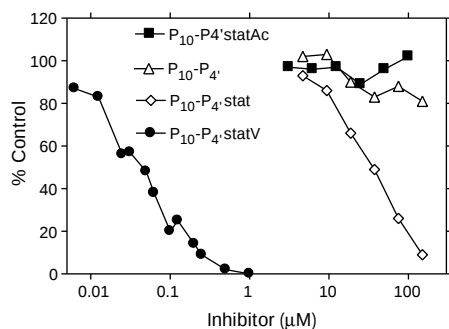


Figure 3 Inhibition of β -secretase by substrate analogue peptides. P₁₀-P₄': Thr662-Phe674 of APP751Swe; stat: statine substituted for P₁ leucine; statAc: substituted statine acetyl ester; statV: substituted statine and valine at P₁'. Activity of β -secretase was assayed as described in Methods, except that peptides were added in DMSO stocks with a final concentration of DMSO of 2%. Values are plotted as percentages of DMSO control.

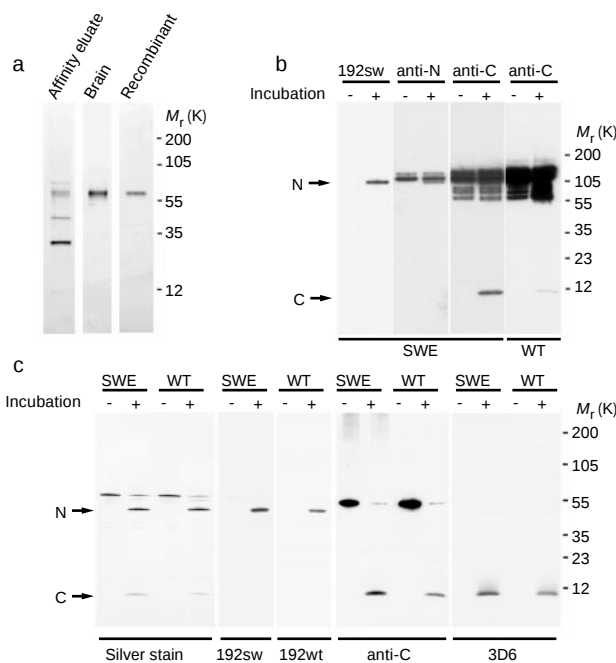


Figure 4 Purification and *in vitro* specificity of β -secretase. **a**, Silver-stained SDS-PAGE gels of β -secretase as eluted from a P₁₀-P₄' StatVal affinity column (affinity) and as finally purified from brain (brain) and β -secretase-transfected 293T cells (recombinant). **b**, Purified full-length APP, wild-type (WT) or Swedish (SWE) and partially purified β -secretase were either incubated (+) or analysed directly (-). Reactions were analysed by immunoblotting: anti-N, antibody 10F1; anti-C, antibody 13G8. Arrows indicate the positions of the N- and C-terminal fragments. **c**, Partially purified β -secretase was incubated with MBP-C125, WT or SWE variants, and analysed by silver-stained SDS-PAGE gels or immunoblotting.

shown), indicating that a single enzyme may be responsible for the human brain β -cleavage activity. In addition to the lack of inhibition by E-64 (cysteine proteinase inhibitor) and leupeptin (serine/cysteine proteinase inhibitor), no inhibition was seen with 3,4-DCIC (serine proteinase inhibitor), iodoacetamide (cysteine proteinase inhibitor) and 1,10-phenanthroline (metalloproteinase inhibitor). Although pepstatin was not inhibitory, even at 50 μ M, some aspartic proteinases are insensitive to pepstatin⁶, and together with the acidic pH preference exhibited by the enzyme, it seemed probable that the β -cleavage was being carried out by a pepstatin-insensitive aspartic proteinase activity.

On the basis of this proposition, a P₁₀-P₄' (ref. 7), P₁ (S)-statine substituted substrate analogue was synthesized. The (S)-statine P₁₀-P₄' analogue dose-dependently inhibited the soluble β -cleavage activity with half-maximal inhibitory concentration (IC₅₀) ~40 μ M, whereas the unmodified peptide (KTEEISEVNLDAEF) was not inhibitory (Fig. 3). Acetylation of the statine hydroxyl, or replacement of the S- with the R-statine enantiomer (data not shown), led to loss of inhibitory activity, indicating that, as in other statine-containing inhibitors for aspartic proteinases⁸, inhibitory potency is dependent on an unmodified hydroxyl residue in the appropriate configuration. Further improvement in the potency was achieved by substituting valine for aspartic acid at the P₁' position, which led to an inhibitor, P₁₀-P₄' StatVal, with IC₅₀ ~30 nM (Fig. 3).

Purification of the β -secretase enzyme activity from human brain was achieved by a sequential four-step procedure, incorporating an affinity-purification step with immobilized P₁₀-P₄' StatVal inhibitor peptide (Fig. 4a, affinity-eluate). The enzyme activity is purified ~300,000-fold using this procedure, yielding a ~70K (relative molecular mass, 70,000) protein homogeneous by silver-stained SDS polyacrylamide gel electrophoresis (PAGE) analysis (Fig. 4a, brain).

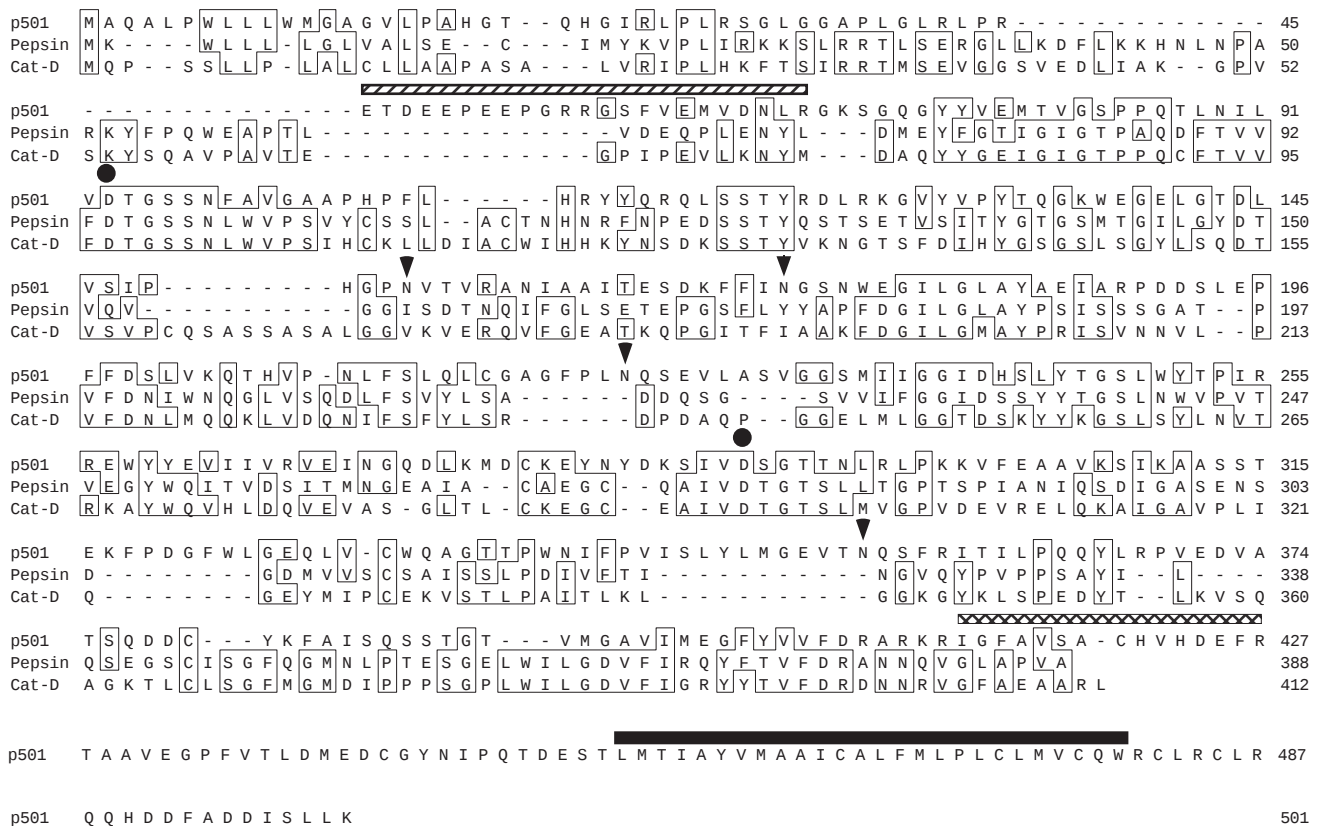


Figure 5 Sequence of β -secretase and homology with known aspartyl proteinases. The predicted amino-acid sequence of β -secretase (p501) is aligned with two well-characterized members of the human aspartyl proteinase family, pepsinogen and cathepsin D. Boxes represent amino acids that are identical in all three sequences. Hatched bar, the N terminus of mature p501, determined by N-terminal sequencing of the

purified protein. Cross-hatched bar, an internal amino-acid sequence from the purified protein determined from a peptidic fragment. Filled bar, transmembrane domain in p501, predicted by hydropathy analysis of the sequence. Filled circles, active-site aspartic acid residues. Arrowheads indicate the four predicted N-linked glycosylation sites in the p501 sequence.

The purified enzyme cleaved both MBPC125Swe and wild-type substrates specifically at their β -cleavage sites (Fig. 4c, silver stain, anti-C), generating N-terminal fragments that were immunoreactive with the β -sAPP neo-epitope-specific Sw192 and Wt192 (ref. 1) antibodies (Fig. 4c, 192sw, 192wt). The corresponding C-terminal fragments (Fig. 4c, 3D6), which reacted with 3D6 (ref. 9), monoclonal antibody specific to the free N terminus of A β 1-5, were also detected, providing evidence of a single endopeptidic cleavage. The cleavage sites were confirmed by N-terminal sequence analyses. Incubation of the brain enzyme with purified full-length APPSwe also resulted in specific cleavage at the β -site, as determined by the generation of single N- and C-terminal fragments (Fig. 4b).

Edman degradation of purified human brain β -secretase revealed a single N-terminal sequence, ETDEEPEEPGRRGSFVEMVDNLR, which we used for isolation of complementary DNA clones encoding full-length β -secretase by a combination of polymerase chain reaction (PCR) and conventional cDNA library screening. Sequence analysis of cDNA clones isolated from human neurons and human foetal brain identified open reading frames that were predicted to encode a polypeptide comprising 501 amino acids, p501. This sequence is identical to a recently described β -secretase clone¹⁰, which was obtained from a 293 cell library using expression cloning. The N-terminal sequence obtained from the purified human brain enzyme corresponded to residues 46–66 of the predicted protein sequence (Fig. 5), indicating that the mature enzyme is generated from a preproenzyme.

Co-transfection of p501 cDNA with either WtAPP751 or SweAPP751 in 293T cells led to marked increases in β -sAPP production (Fig. 6b) and substantial increases in A β peptide (Fig. 6a). The large increase in β -sAPP is accompanied by a corresponding

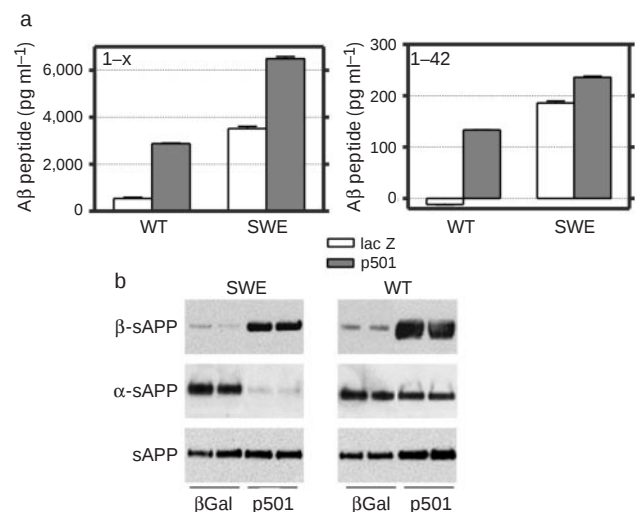


Figure 6 Transfection of cells with β -secretase. Cells were co-transfected with either wild-type (WT) or Swedish (SWE) APP and either the β -secretase clone p501 or β -galactosidase control. Co-expression of β -gal or p501 and APP coding sequences was directed from the CMV promoter peptide in vectors that did not permit replication in 293T cells. **a**, Media were analysed for A β peptide using ELISAs specific for all A β C termini (1–x) or A β 42 (1–42). **b**, Media were analysed by immunoblotting to measure secreted APP (sAPP): α -sAPP (antibody 1736), β -sAPP (antibodies 192 or 192Swe) and total sAPP (antibody 8E5).

decrease in α -sAPP levels (Fig. 6b, α -sAPP), suggesting a very efficient cleavage of full-length APP at the β -cleavage site. Thus, p501 overexpression leads to increased β -secretase cleavage of both WtAPP and SweAPP, accompanied by increased concentrations of A β peptide.

The recombinant enzyme was purified using cation exchange and the inhibitor affinity matrix (Fig. 4a, recombinant), and its N-terminal sequence was determined. About 90% of the purified protein N-terminal sequence corresponded exactly to that determined for the purified brain enzyme (residue 46). About 10% of the total sequence started at residue 22, the proenzyme form generated upon removal of the signal peptide (residues 1–21). Expression of the recombinant enzyme is thus accompanied by efficient cellular proteolytic processing indistinguishable from that of the native enzyme purified from human brain.

Northern blot analyses (data not shown) showed high expression in brain and pancreas, as previously described¹⁰; however, whereas brain displayed high enzymatic activity (Fig. 2a), pancreas had very low activity. It may be that post-transcriptional regulation, possibly at the level of translation, regulates pancreas β -secretase activity.

The activity of human brain β -secretase, both *in vitro* and in cells, is highly specific for the β -cleavage site. Overexpression of the enzyme leads to marked increases in β -sAPP and A β peptide levels. The enzyme shows enhanced cleavage of Swedish as compared to wild-type substrate, in agreement with results obtained from cellular studies. The highest levels of enzyme activity are found in cells and tissues of CNS origin, which supports its proposed role as human brain β -secretase. The pH-activity profile of the enzyme is acidic, consistent with cellular studies of APP metabolism that intracellular acidic compartments are involved in β -, but not α -, secretase cleavage events⁴.

Our data show that activation of β -secretase from the proenzyme form occurs efficiently in cell lines. β -secretase enzyme activity levels correlate well with the propensity of the cell line or tissue to generate A β peptide from APP. As with HIV proteinase¹¹, and illustrated here by the peptide-based transition state analogue, potent and selective inhibitors of this enzyme can be generated. Further optimization with structure-based drug design should lead to the discovery of small molecules that would be effective in inhibiting A β peptide production in Alzheimer's disease. □

Methods

APP substrates

The 125 C-terminal residues of APP, in both the wild-type and the Swedish variants, were fused to the C terminus of MBP in the *Escherichia coli* expression vector pMALc (New England Biolabs). The fusion proteins were induced in bacterial cultures and purified by solubilization with 1% Triton X-100/7 M urea, followed by chromatography on amylose-agarose. The bound proteins were eluted with 10 mM maltose in 20 mM Tris, pH 7.5, 1% Triton X-100, 0.1 M NaCl, diluted 1:1 with 6 M guanidine, and stored in frozen aliquots (0.1 mg ml⁻¹ fusion protein in 10 mM Tris, pH 7.5, 0.2% Triton X-100, 0.15 M guanidine-HCl).

Full-length APP was prepared from 293 cells stably transfected with Swedish or wild-type APP. P2 membranes were washed with 20TE (20 mM Tris, 2 mM EDTA) with 0.5 M NaCl and extracted with 20TE with 0.5% Triton X-100. Full-length APP was purified by anion exchange and immuno-affinity chromatography¹² on 13G8 coupled to NHS-Sepharose.

β -secretase enzyme assay

β -cleavage assays were carried out in 20 mM sodium acetate, pH 4.8, 0.06% Triton X-100, with 10 μ g ml⁻¹ MBPAPP125. Reaction mixtures were incubated at 37°C for 1–2 h, and the quenched reaction mixtures were then loaded onto 96-well plates coated with a polyclonal antibody raised to MBP. Generated β -cleaved product was detected using biotinylated Sw192 or biotinylated Wt192 as specific reporter antibodies and quantitated against the appropriate MBP-C26 standard.

Enzyme isolation

Crude membrane fractions were obtained from frozen human brain tissue. The membrane pellets were extracted with 20 mM MES, pH 6, 0.15 M NaCl, 0.5% Triton X-100. The soluble extract obtained by centrifugation was loaded onto WGA-agarose at pH 7.5, and the column eluted with 6% chitin hydrolysate after washing. We diluted the eluate fourfold into 20 mM sodium acetate, pH 5, adjusted the pH to 5.0, and passed it through a Hi-Trap SP column. The SP flow-through was adjusted to pH 4.5 with acetic acid and applied to a column of P₁₀-P₄ statV coupled to NHS-Sepharose. The column was washed and eluted

with 0.2 M NaCl, 40 mM sodium borate, pH 9.5, and 2.0% Triton X-100. We removed contaminating proteins by anion exchange on MiniQ. N-terminal sequence analysis was carried out by Argo Bioanalytica.

Cloning and expression

Screening of the Origene human foetal brain Rapid-Screen cDNA Library Panel by degenerate PCR, using oligos 5'-GAGAGACGA(G/A)GA(G/A)CC(A/T)GAGGAGCC-3' and 5'-CGTCACAG(G/A)TT(G/A)TCTACCATCTC-3', identified a partial clone encoding the amino-acid sequence from the purified enzyme. Radiolabelled probe derived from this clone was used to screen a size-selected human foetal primary neuronal cell library generated in a mammalian expression vector to obtain the full-length clone, p501. Wild-type or Swedish APP751 was transiently co-transfected with either p501 or β -gal control into 293T cells (Edge Biosciences). Co-expression of β -gal or p501 and APP coding sequences was directed from the CMV promoter in vectors that did not permit replication in 293T cells. Medium was collected 48 h after transfection.

Antibodies

Antibodies were directed against the following portions of the APP751 sequence: 10F1, residues 20–304 (ref. 9); 13G8, 727–751; 8E5, 495–643 (ref. 9); 192wt and 192sw, specific for β -cleaved APP wild-type and Swedish variants, respectively¹⁴; 1746, specific for α -cleaved APP (courtesy of D. Selkoe); 3D6, A β 1–5, specific for the free N terminus⁵.

Received 1 November; accepted 8 November 1999.

- Seubert, P. *et al.* Secretion of beta-amyloid precursor protein cleaved at the amino terminus of the beta-amyloid peptide. *Nature* **361**, 260–263 (1993).
- Mullan, M. *et al.* A pathogenic mutation for probably Alzheimer's disease in the APP gene at the N-terminus of beta-amyloid. *Nature Genet.* **1**, 345–347 (1992).
- Citron, M. *et al.* Mutation of the beta-amyloid precursor protein in familial Alzheimer's disease increases beta-protein secretion. *Nature* **360**, 672–674 (1992).
- Knops, J. *et al.* Cell-type and amyloid protein-type specific inhibition of Ab release by bafilomycin A1, a selective inhibitor of vacuolar ATPases. *J. Biol. Chem.* **270**, 2419–2422 (1995).
- Valdes, F., Munoz, C., Faria-Velasco, A. & Orrego, F. Subcellular distribution of rat brain cortex high-affinity, sodium-dependent, glycine transport sites. *Brain Res.* **122**, 95–112 (1977).
- Narutaki, S., Dunn, B. M. & Oda, K. Subsite preference of pepstatin-insensitive carboxyl proteinases from bacteria. *J. Biochem. (Tokyo)* **125**, 75–81 (1999).
- Schechter, I. & Berger, A. On the size of the active site in proteinases. I. Papain. *Biochem. Biophys. Res. Commun.* **27**, 157–162 (1967).
- Rich, D. H. in *Proteinase Inhibitors* (eds Barrett, A. J. & Salvesen, G.) (Elsevier Science, Amsterdam, 1986).
- Johnson-Wood, K. *et al.* Amyloid precursor protein processing and A β ₄₂ deposition in a transgenic mouse model of Alzheimer disease. *Proc. Natl Acad. Sci. USA* **94**, 1550–1555 (1997).
- Vassar, R. *et al.* β -secretase cleavage of the Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE. *Science* **286**, 735–741 (1999).
- Lin, J. H., Ostovic, D. & Vacca, J. P. The integration of medicinal chemistry, drug metabolism, and pharmaceutical research and development. The story of Crixivan, an HIV proteinase inhibitor. *Pharm. Biotechnol.* **11**, 233–255 (1998).
- Knops, J. *et al.* Isolation of Baculovirus-derived secreted and full-length β -amyloid precursor protein. *J. Biol. Chem.* **266**, 7285–7290 (1991).

Correspondence and requests for materials should be addressed to S.S. (e-mail: ssinha@elanpharma.com). The nucleotide sequence of p501 has been deposited in the GenBank under accession number AF201468.

Axon routing across the midline controlled by the *Drosophila* Derailed receptor

Joshua L. Bonkowsky*, Shingo Yoshikawa*, David D. O'Keefe, Audra L. Scully & John B. Thomas

Molecular Neurobiology Laboratory, The Salk Institute for Biological Studies, P.O. Box 85800, San Diego, California 92186, USA

* These authors contributed equally to this work

In nervous systems with symmetry about the midline, many neurons project axons from one side to the other. Although several of the components controlling midline crossing have been identified^{1–4}, little is known about how axons choose the appropriate pathway when crossing. For example, in the *Drosophila* embryo axons cross the midline in one of two distinct tracts, the anterior or posterior commissure (AC or PC, respec-

tively). Here we show that the Derailed (Drl) receptor tyrosine kinase is expressed by neurons that project in the AC, and that in the absence of Drl such neurons often project abnormally into the PC. Conversely, misexpression of Drl in PC neurons forces them to cross in the AC. The behaviour of Drl-misexpressing neurons and the *in vivo* binding pattern of a soluble Drl receptor probe indicate that Drl acts as a guidance receptor for a repellent ligand present in the PC. Our results show that Drl is a novel component in the control of midline crossing.

Within the *Drosophila* embryonic ventral nerve cord (VNC), the Drl receptor is expressed exclusively by neurons that project in the AC (Fig. 1a–c). These include locally projecting interneurons, intersegmental interneurons and a small number of motor neurons^{5,6}. Drl is localized to the growth cones and axons of neurons as they extend into and through the AC, but is rapidly down-regulated when they leave. This highly restricted expression implies a role for the Drl receptor in axon guidance through the AC.

To examine the commissures of *drl* mutants, we used an antibody that labels all axons. We found that, in embryos mutant for the null allele *drl*^{R343}, 30% (*n* = 170) of segments within the VNC contained axons that projected between the AC and PC (Fig. 1e), a behaviour not seen in the wild type (Fig. 1d). To show that the commissural defects in the VNC arise from pathfinding errors and not from alterations in glial or neuronal lineages, we examined expression of a number of neuronal and glial markers including Islet, Lim3, Sim and AA142 (refs 7–10). The expression of these markers in *drl* mutants was indistinguishable from that of the wild type. The midline glial cells, which are essential for the formation of the commissures⁹, appeared morphologically normal.

We could not distinguish between pathfinding errors made by AC or PC axons using pan-axonal labelling. Therefore, we used a set of *tau-lacZ* marker lines¹¹ to examine the axonal projections of small subsets of AC and PC axons. To visualize AC axons we used *drl*^{Red2}, a null allele of *drl* retaining a *tau-lacZ* P-element which expresses the axon-targeted reporter Tau-β-galactosidase (Tau-β-gal) in a subset of the Drl neurons. In *drl*^{Red2/+} heterozygotes, the labelled AC neurons projected normally and never entered the PC (*n* = 120) (Fig. 1f). However, in *drl*^{Red2/drl} embryos, 34% of segments had AC axons that projected aberrantly in the PC (*n* = 127) (Fig. 1g). These segments contained axons that projected entirely through the PC, as well as axons that left the AC and entered the PC near the midline. In contrast to the abnormal behaviour of AC axons, PC axons in *drl* mutants did not appear to cross into the AC as assayed with *tau-lacZ* lines 3.438 or 3.538. Thus, the observed commissural defects in *drl* mutants arise from abnormal crossings of AC axons into the PC. As we do not have specific markers for all of the PC neurons, we cannot rule out the possibility that some PC axons might also project aberrantly in *drl* mutants, perhaps following the abnormally projecting AC axons.

To test whether the Drl receptor is sufficient to instruct axons to project in the AC, we ectopically expressed it in PC neurons using the GAL4/UAS transactivation system¹². For these experiments we used *eagle-GAL4* (*eg-GAL4*) (refs 13, 14), a driver which expresses GAL4 in a sufficiently small subset of PC neurons that we could unambiguously follow their axonal projections. As assayed with a *UAS-tau-lacZ* reporter transgene, *eg-GAL4* drives expression in two small clusters of Eg-expressing interneurons in each hemisegment, both of which project axons across the midline (Fig. 2a). One cluster projects in the PC and the other in the AC. At the midline, the axons from homologous clusters on either side of the VNC fasciculate with one another, forming two distinct bundles of axons, one within each of the commissures. When forced to misexpress Drl using two copies of a *UAS-drl* transgene, the Eg neurons in the AC were unaffected, indicating that overexpression of Drl in neurons that normally express it has no effect on their pathfinding. However, the Eg neurons that normally cross in the PC switched their projections to the AC in all segments (*n* = 169) (Fig. 2b; Table 1).

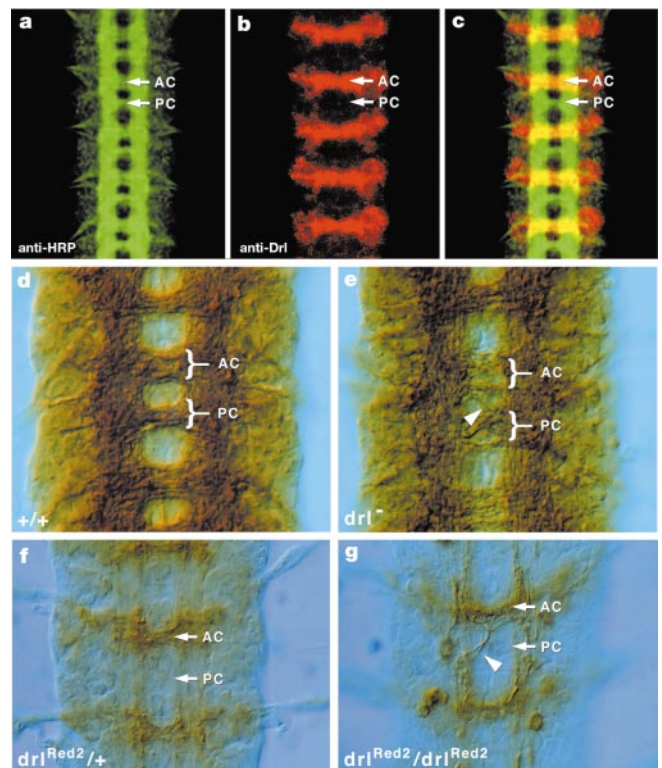


Figure 1 Drl is required for proper projection of anterior commissure axons across the midline. **a–c**, Double-labelling of a wild-type embryonic VNC with anti-HRP (green), which labels all axons, and anti-Drl (red). The merge is shown in **c**. In each segment, axons traverse the midline in one of two axon tracts, the anterior or posterior commissure (AC and PC). Drl is expressed by neurons that project axons in the AC. Drl is downregulated on these axons when they leave the AC. **d, e**, Wild-type (**d**) and *drl*^{R343} mutant (**e**) embryos stained with monoclonal antibody BP102 which labels all axons. In *drl* mutants, many segments show axons abnormally crossing between the AC and PC (arrowhead). **f, g**, Anti-β-gal stained embryos carrying *drl*^{Red2}, a null allele of *drl* that expresses Tau-β-gal in a subset of AC neurons. In *drl*^{Red2/+} (**f**) all stained axons project normally in the AC. In *drl*^{Red2/drl} mutants (**g**) AC axons project abnormally between the AC and PC (arrowhead).

In 98% of segments, every Eg-expressing PC axon switched to the AC, whereas in 2% of segments one or a few axons remained in the PC. Decreasing the levels of Drl misexpression either by using a single copy of *UAS-drl* or by lowering the temperature resulted in fewer PC-to-AC switched axons (Table 1). Thus, Drl's capacity to switch PC projections to the AC is dose-dependent. The ability of Drl to switch projections also depends upon an intact extracellular domain, as deletion of 155 amino acids of the extracellular domain (DrlΔE) rendered it incapable of switching any of the Eg PC axons when expressed from two copies of *UAS-drl*ΔE (Fig. 2c; Table 1).

Despite the fact that the Drl-misexpressing Eg PC neurons had switched their projections to the AC, they still retained key aspects of their original identity. As they would normally do in wild-type embryos, these neurons recognized and fasciculated specifically with their contralateral homologues, forming an axon bundle within the AC distinct from the normal Eg AC bundle (Fig. 2b). In addition, the switched axonal projections remain locally within the segment, just as they would in the wild type.

Next we asked whether Drl was sufficient to cause non-crossing neurons to enter the AC. For this experiment we used *apterous*^{GAL4} (*ap*^{GAL4}), which drives expression specifically in the Ap neurons¹⁵, a small subset of intersegmental interneurons which do not express Drl and project axons ipsilaterally (Fig. 3a). When forced to misexpress Drl, the Ap neurons remained ipsilateral and did not enter the AC, indicating that Drl is not sufficient to guide these neurons into the AC. However, when the Ap neurons were forced to cross the midline by misexpression of Commissureless (Comm), a com-

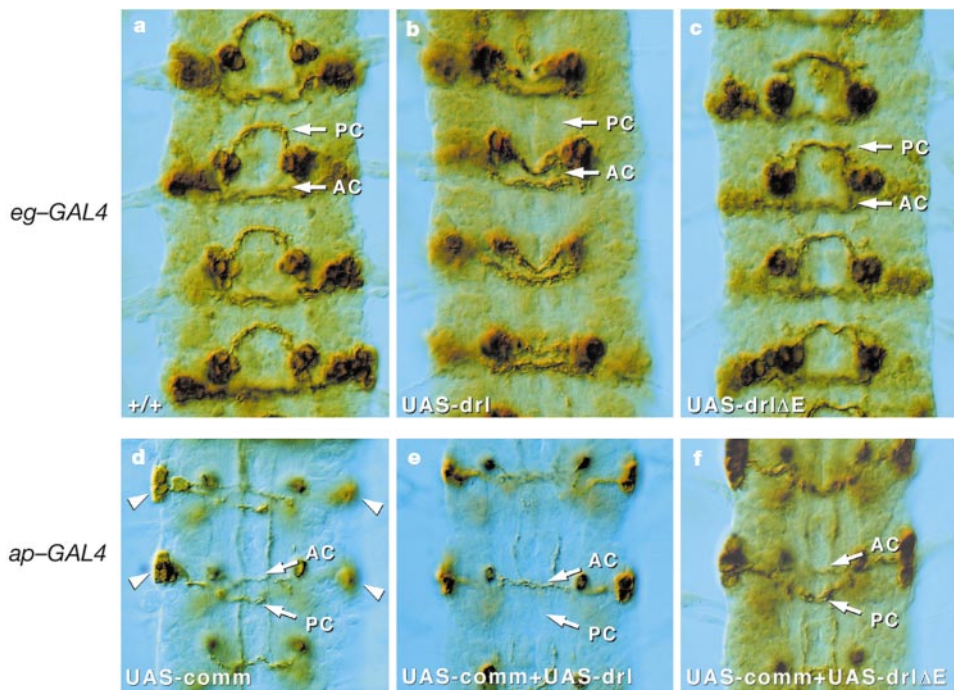


Figure 2 Drl switches axons from the PC to the AC. **a–c**, Embryos carrying the *eg-GAL4* driver plus *UAS* transgenes. In the wild type (**a**), two clusters of Eg neurons extend axons across the midline, as assayed by staining for the axon-targeted Tau-β-gal reporter expressed from a *UAS-tau-lacZ* transgene. Each cluster of neurons forms a distinct axon bundle, one in the AC and one in the PC. When forced to misexpress Drl from two copies of a *UAS-drl* transgene (**b**), all of the Eg PC axons cross in the AC, forming a bundle distinct from the normal Eg AC bundle. Misexpression of DrlΔE from two copies of a *UAS-drlΔE* transgene (**c**) has no effect on the Eg PC axons. **d–f**, Embryos carrying the *ap^{GAL4}* driver plus *UAS* transgenes. When misexpressing Comm from a *UAS-comm* transgene (**d**), the

thoracic-specific lateral cluster of Ap neurons (arrowheads) project axons in the PC, whereas the other Ap neurons project axons in the AC. When forced to co-misexpress Drl and Comm (**e**), the thoracic-specific Ap axons switch from the PC to the AC. When co-expressed with Comm, DrlΔE cannot switch projections (**f**). Genotypes: *eg-GAL4/UAS-tau-lacZ* (**a**), *UAS-drl₁ UAS-drl₂* +; *eg-GAL4/UAS-tau-lacZ* (**b**), *UAS-tau-lacZ* +; *eg-GAL4/UAS-drlΔE₁ UAS-drlΔE₂* (**c**), *ap^{GAL4}* +; *UAS-comm UAS-tau-lacZ* + (**d**), *ap^{GAL4}* +; *UAS-comm UAS-tau-lacZ* +; *UAS-drl₁ UAS-drl₂* + (**e**), *ap^{GAL4}* +; *UAS-comm UAS-tau-lacZ/UAS-drlΔE₁ UAS-drlΔE₂* (**f**).

ponent of the midline-crossing mechanism^{3,16}, Drl dictated which commissure they chose. Misexpression of Comm caused all Ap neurons to cross the midline. Thoracic-specific Ap neurons misexpressing Comm crossed in the PC, whereas the other Comm-misexpressing Ap neurons crossed in the AC (Fig. 2d). Co-misexpression of Drl switched the projections of all thoracic-specific Ap neurons from the PC to the AC in 100% of the segments examined (*n* = 30; Fig. 2e), whereas DrlΔE was incapable of switching any axons to the AC (*n* = 39) (Fig. 2f).

From the above experiments, we conclude that Drl instructs neurons that cross the midline to choose the AC. To address whether Drl-expressing axons choose the AC because of an attractive activity in the AC or a repulsive one in the PC, we examined in more detail the effects of Drl misexpression by the Ap neurons without co-misexpression of Comm. In the wild type, the growth cones of the Ap interneurons choose a single medial ipsilateral pathway within each connective¹⁷. Upon reaching the adjacent anterior segment they tightly fasciculate with their homologues, forming a discrete axon bundle running the length of the VNC (Fig. 3a). In the course of their normal projection, the Ap neurons encounter the lateral edges of both the AC and PC. Misexpression of Drl in the Ap neurons caused their growth cones to stall as they approached the PC (Fig. 3b), suggesting that a repulsive ligand in the PC is recognized by the Drl receptor. Such a repulsive ligand in the PC would also explain why Drl-misexpressing Eg PC neurons project in the AC of the adjacent posterior segment rather than projecting across the edge of the PC to the AC of their segment of origin (Fig. 2b).

To address further the idea that Drl functions as a guidance receptor for a repellant ligand in the PC we capitalized on the finding that soluble epitope-tagged extracellular domains of many receptors can be used as probes *in vivo* to determine the spatial distribution of

their ligands^{18,19}. We fused the entire extracellular domain of Drl to five copies of the c-myc epitope (creating Drl-myc). In addition, we constructed a control probe (DrlΔE-myc) which lacked the same portion of the extracellular domain as DrlΔE. Both secreted proteins had the expected sizes by western blot analysis. Staining of embryos revealed a specific binding pattern in the VNC for Drl-myc; identical patterns were obtained regardless of whether the epitope tag used was c-myc, alkaline phosphatase or the Fc fragment of the human immunoglobulin molecule (M. Kokel and J.B.T., unpublished data). In contrast, no staining pattern was detected using the DrlΔE-myc control probe, indicating that Drl binding, as well as its ability to switch commissural axons between commissures, is dependent upon an intact extracellular domain.

Drl-myc binding was restricted to the PC and to globular domains at the intersection of the PC and longitudinal connectives (Fig. 3c), a pattern complementary to the AC-restricted expression of Drl. As assayed by double-labelling of *drl^{P3.765}* embryos, which express the axon-targeted reporter Tau-β-gal in the Drl neurons⁵, the Drl axons clearly avoid the zones of Drl-myc binding as they

Table 1 Dose-dependence of PC-to-AC axon switching by Drl

	Percent of segments		
	All axons switched	Some axons switched	No axons switched
Wild type	0	0	100 (<i>n</i> = 100)
<i>UAS-drl</i> 2 copies 29°C	98	2	0 (<i>n</i> = 169)
<i>UAS-drl</i> 1 copy 29°C	75	18	7 (<i>n</i> = 96)
<i>UAS-drl</i> 1 copy 25°C	55	37	8 (<i>n</i> = 98)
<i>UAS-drlΔE</i> 2 copies 29°C	0	0	100 (<i>n</i> = 100)

In addition to the listed *drl* transgenes, all embryos carried the *eagle-GAL4* driver plus a *UAS-tau-lacZ* transgene to label axons.

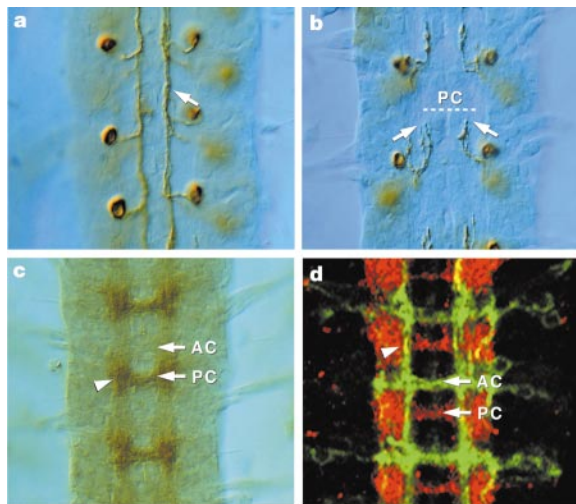


Figure 3 Drl recognizes a repulsive cue in the PC. **a, b**, Behaviour of the Ap neurons as assayed using ap^{GAL4} driving expression of the axon-targeted reporter Tau- β -gal from a UAS -tau- $lacZ$ transgene. In the wild-type (**a**) the Ap neurons project anteriorly in a medial pathway within the ipsilateral connective, traversing the lateral edges of both the AC and PC. When forced to misexpress Drl from two copies of a UAS - drl transgene (**b**) the growth cones of the Ap neurons (arrows) stall at the level of the PC (dashed line marks the posterior boundary of the PC). Genotypes: $ap^{GAL4}/+$; UAS -tau- $lacZ$ +/+ (**a**), ap^{GAL4}/UAS - drl_1 UAS - drl_2 ; UAS -tau- $lacZ$ +/+ (**b**). **c**, A wild-type embryo incubated with the Drl-myc soluble extracellular domain probe. Drl-myc binding is restricted to the PC and to globular domains at the junction of the PC and the two connectives (arrowhead), a pattern complementary to that of the Drl receptor, which is restricted to the AC. **d**, A $drl^{P3.765}/+$ embryo double-labelled for Drl-myc (red) and Tau- β -gal (green), which is expressed in the Drl neurons. As the Drl neurons project into the AC, they avoid regions of Drl-myc binding. After leaving the AC, the Drl axons do traverse a region of Drl-myc binding in the connective (arrowhead), but they have downregulated Drl on their growth cones and axons by this time (see Fig. 1b).

project into and through the AC (Fig. 3d). Within the connectives, the axons of Drl intersegmental interneurons do traverse a region of putative ligand expression. However, Drl is downregulated on the growth cones and axons of these neurons when they leave the AC and begin extending in the connectives (Fig. 1b, c). Thus, it appears that the downregulation of the Drl repulsive guidance receptor allows the intersegmental growth cones to traverse this area, a model similar to the downregulation of Robo, the receptor for the midline repellent Slit, allowing growth cones to cross the midline²⁻⁴.

We propose that crossing the midline involves at least three distinct mechanisms. Growing axons are attracted to the midline by diffusible cues such as the Netrins²⁰⁻²⁴, their crossing is controlled by the Slit/Robo/Comm guidance mechanism^{2-4,16,25} and their choice of commissure is dictated by the Drl guidance mechanism. The fact that in drl mutants many axons still project appropriately in the AC indicates that Drl may be part of a multicomponent system that ensures proper commissure choice. For example, it seems likely that there are additional attractive cues for the AC and that, once neurons are attracted there, Drl ensures that none enters the PC. Drl may also have a similar function in commissure choice within the adult nervous system in *Drosophila*, where it appears to be required for appropriate projections between the mushroom bodies^{26,27}. In drl mutants, the Drl intersegmental neurons project abnormally in the connectives after crossing the midline, despite the fact that by this time Drl has been downregulated on their growth cones and axons⁵. It thus appears that proper crossing of the midline is a prerequisite for subsequent guidance events in the connectives. Defects in these subsequent guidance events could arise as a consequence of the axons crossing in the wrong commissure or of failure to receive the Drl signal itself during crossing. Finally, given the existence of vertebrate Drl homologues²⁸ and the apparent evolutionary conservation of other

midline crossing components, it is possible that a Drl-like guidance mechanism operates in all complex nervous systems. □

Methods

Fly strains and genetics

P-element transformation was carried out using standard methods²⁹. For independent genotyping of embryos, lines were balanced over the marked chromosomes CyO , $wingless$ - $lacZ$ or $TM3$, $Actin$ - $lacZ$.

Soluble Drl receptor probe

The extracellular domain (nucleotides 327–1053) of the drl complementary DNA was amplified by polymerase chain reaction and cloned as a *Bam*HI fragment into pFB-myc, a modified pFastBac vector (Gibco-BRL) containing five copies of the *c-myc* epitope cloned into the *Bam*HI–*Eco*RI site, to create pFB-Drl-myc. pFB-Drl Δ E-myc was created by excising the *Nar*I fragment from pFB-Drl-myc and blunt-end ligating. Baculovirus protein expression was performed as described for the Bac-to-Bac system (Gibco BRL). For maximum protein expression, Hi5 cells were infected at ~70% confluence, and the medium was collected after six days and stored in 0.1% sodium azide at 4 °C.

Immunohistochemistry

All immunohistochemistry was performed as described³⁰. For staining with Drl-myc, dissected embryos were incubated in Drl-myc-containing medium for 30 min at room temperature and washed three times for 5 min each in PBS, followed by fixation and immunohistochemistry with anti-myc antibodies. $drl^{P3.765}$ embryos were double-labelled for myc and β -gal expression by stepwise incubation with anti-myc, biotin-conjugated anti-mouse and Cy-3-conjugated streptavidin (Jackson Laboratories); embryos were then fixed for 10 min in 4% paraformaldehyde and incubated stepwise with anti- β -gal and FITC-conjugated anti-rabbit (Vector Laboratories). Fluorescently labelled preparations were imaged using a Zeiss LSM 510 confocal microscope and images compiled using Adobe Photoshop.

Received 15 July; accepted 16 September 1999.

1. Brose, K. *et al.* Slit proteins bind Robo receptors and have an evolutionarily conserved role in repulsive axon guidance. *Cell* **96**, 795–806 (1999).
2. Kidd, T. *et al.* Roundabout controls axon crossing of the CNS midline and defines a novel subfamily of evolutionarily conserved guidance receptors. *Cell* **92**, 205–215 (1998).
3. Kidd, T., Russell, C., Goodman, C. S. & Tear, G. Dosage-sensitive and complementary functions of roundabout and commissureless control axon crossing of the CNS midline. *Neuron* **20**, 25–33 (1998).
4. Kidd, T., Bland, K. S. & Goodman, C. S. Slit is the midline repellent for the robo receptor in *Drosophila*. *Cell* **96**, 785–794 (1999).
5. Callahan, C. A., Muralidhar, M. G., Lundgren, S. E., Scully, A. L. & Thomas, J. B. Control of neuronal pathway selection by a *Drosophila* receptor protein-tyrosine kinase family member. *Nature* **376**, 171–174 (1995).
6. Bonkowski, J. L. & Thomas, J. B. Cell-type specific modular regulation of *derailed* in the *Drosophila* nervous system. *Mech. Dev.* **82**, 181–184 (1999).
7. Crews, S. T., Thomas, J. B. & Goodman, C. S. The *Drosophila* single-minded gene encodes a nuclear protein with sequence similarity to the *per* gene product. *Cell* **52**, 143–151 (1988).
8. Klämbt, C., Jacobs, J. R. & Goodman, C. S. The midline of the *Drosophila* central nervous system: a model for the genetic analysis of cell fate, cell migration, and growth cone guidance. *Cell* **64**, 801–815 (1991).
9. Thor, S. & Thomas, J. B. The *Drosophila* islet gene governs axon pathfinding and neurotransmitter identity. *Neuron* **18**, 397–409 (1997).
10. Thor, S., Andersson, S. G., Tomlinson, A. & Thomas, J. B. A LIM-homeodomain combinatorial code for motor-neuron pathway selection. *Nature* **397**, 76–80 (1999).
11. Callahan, C. A. & Thomas, J. B. Tau- β -galactosidase, an axon-targeted fusion protein. *Proc. Natl Acad. Sci. USA* **91**, 5972–5976 (1994).
12. Brand, A. H. & Perrimon, N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401–415 (1993).
13. Dittich, R., Bossing, T., Gould, A. P., Technau, G. M. & Urban, J. The differentiation of the serotonergic neurons in the *Drosophila* ventral nerve cord depends on the combined function of the zinc finger proteins Eagle and Hucklebein. *Development* **124**, 2515–2525 (1997).
14. Higashijima, S., Shishido, E., Matsuzaki, M. & Saigo, K. *eagle*, a member of the steroid receptor gene superfamily, is expressed in a subset of neuroblasts and regulates the fate of their putative progeny in the *Drosophila* CNS. *Development* **122**, 527–536 (1996).
15. O'Keefe, D. D., Thor, S. & Thomas, J. B. Function and specificity of LIM domains in *Drosophila* nervous system and wing development. *Development* **125**, 3915–3923 (1998).
16. Tear, G. *et al.* *commissureless* controls growth cone guidance across the CNS midline in *Drosophila* and encodes a novel membrane protein. *Neuron* **16**, 501–514 (1996).
17. Lundgren, S. E., Callahan, C. A., Thor, S. & Thomas, J. B. Control of neuronal pathway selection by the *Drosophila* LIM homeodomain gene *apterous*. *Development* **121**, 1769–1773 (1995).
18. Cheng, H.-J. & Flanagan, J. G. Identification and cloning of ELF-1, a developmentally expressed ligand for the Mek4 and Sek receptor tyrosine kinases. *Cell* **79**, 157–168 (1994).
19. Cheng, H.-J., Nakamoto, M., Bergemann, A. D. & Flanagan, J. G. Complementary gradients in expression and binding of ELF-1 and Mek4 in development of the topographic retinotectal projection map. *Cell* **82**, 371–381 (1995).
20. Kolodziej, P. A. *et al.* *frazzled* encodes a *Drosophila* member of the DCC immunoglobulin subfamily and is required for CNS and motor axon guidance. *Cell* **87**, 197–204 (1996).
21. Kennedy, T. E., Serafini, T., de la Torre, J. R. & Tessier-Lavigne, M. Netrins are diffusible chemotropic factors for commissural axons in the embryonic spinal cord. *Cell* **78**, 425–435 (1994).
22. Mitchell, K. J. *et al.* Genetic analysis of Netrin genes in *Drosophila*: Netrins guide CNS commissural axons and peripheral motor axons. *Neuron* **17**, 203–215 (1996).

23. Harris, R., Sabatelli, L. M. & Seeger, M. A. Guidance cues at the *Drosophila* CNS midline: identification and characterization of two *Drosophila* Netrin/UNC-6 homologs. *Neuron* **17**, 217–228 (1996).
24. Chan, S. S. *et al.* UNC-40, a *C. elegans* homolog of DCC (Deleted in Colorectal Cancer), is required in motile cells responding to UNC-6 netrin cues. *Cell* **87**, 187–195 (1996).
25. Seeger, M., Tear, G., Ferres, M. D. & Goodman, C. S. Mutations affecting growth cone guidance in *Drosophila*: genes necessary for guidance toward or away from the midline. *Neuron* **10**, 409–426 (1993).
26. Moreau-Fauvarque, C., Taillebourg, E., Boissoneau, E., Mesnard, J. & Dura, J. M. The receptor tyrosine kinase gene *linotte* is required for neuronal pathway selection in the *Drosophila* mushroom bodies. *Mech. Dev.* **78**, 47–61 (1998).
27. Simon, A. F., Boquet, I., Synguelakis, M. & Preat, T. The *Drosophila* putative kinase *linotte* (derailed) prevents central brain axons from converging on a newly described interhemispheric ring. *Mech. Dev.* **76**, 45–55 (1998).
28. Hovens, C. M. *et al.* RYK, a receptor tyrosine kinase-related molecule with unusual kinase domain motifs. *Proc. Natl Acad. Sci. USA* **89**, 11818–11822 (1992).
29. Rubin, G. M. & Spradling, A. C. Genetic transformation of *Drosophila* with transposable element vectors. *Science* **218**, 348–353 (1982).
30. Callahan, C. A., Bonkovsky, J. L., Scully, A. L. & Thomas, J. B. *derailed* is required for muscle attachment site selection in *Drosophila*. *Development* **122**, 2761–2767 (1996).

Acknowledgements

We thank T. Ishihara, G. Technau, C. Goodman, C. Klämbt and the Indiana *Drosophila* Stock Center for stocks, and G. Lemke and members of the Thomas lab for helpful comments. This work was supported by funds from the UC San Diego MSTP program to J.L.B., an HFSP Long-Term Fellowship to S.Y. and grants from the NIH to J.B.T.

Correspondence and requests for materials should be addressed to J.B.T. (e-mail: jthomas@salk.edu).

Bazooka provides an apical cue for Inscuteable localization in *Drosophila* neuroblasts

Andreas Wodarz, Andreas Ramrath, Ute Kuchinke & Elisabeth Knust

Institut für Genetik, Heinrich-Heine-Universität Düsseldorf, Universitätsstrasse 1, 40225 Düsseldorf, Germany

Asymmetric cell division generates daughter cells with different developmental fates from progenitor cells that contain localized determinants. During this division, the asymmetric localization of cell-fate determinants and the orientation of the mitotic spindle must be precisely coordinated. In *Drosophila* neuroblasts, *inscuteable* controls both spindle orientation and the asymmetric localization of the cell-fate determinants Prospero and Numb¹. Inscuteable itself is localized in an apical cortical crescent and thus reflects the intrinsic asymmetry of the neuroblast^{1,2}. Here we show that localization of Inscuteable depends on Bazooka, a protein containing three PDZ domains with overall sequence similarity to Par-3 of *Caenorhabditis elegans*³. Bazooka and Inscuteable form a complex that also contains Stauf, a protein responsible for the asymmetric localization of *prospero* messenger RNA. We propose that, after delamination of the neuroblast from the neuroepithelium, Bazooka provides an asymmetric cue in the apical cytocortex that is required to anchor Inscuteable. As Bazooka is also responsible for the maintenance of apical–basal polarity in epithelial tissues⁴, it may be the missing link between epithelial polarity and neuroblast polarity.

Drosophila neuroblasts divide asymmetrically, giving rise to another neuroblast and a smaller, basally located ganglion mother cell (GMC). In these divisions, the cell-fate determinant Prospero (Pros) is segregated into the GMC only, where it activates transcription of GMC-specific genes and suppresses neuroblast-specific gene expression^{5–7}. Segregation of Pros into the GMC is controlled by a cascade of genes, which are ordered in a functional hierarchy^{8,9}. A key protein in this process is Inscuteable (Insc), which has five ankyrin-like domains and controls both spindle orientation and the

asymmetric localization of particular proteins and mRNAs^{1,2}. Insc is localized in the apical cytocortex of neuroblasts^{1,2}, where it binds directly to Miranda (Mir) and Stauf (Stau)^{10,11}. Mir, an adaptor protein, binds to Pros, Numb and Stau; Stau itself is required to localize *pros*-mRNA^{10,12–16}. In *insc* mutants, Mir and Stau are mis-localized. Conversely, apical localization of Insc is normal in mutants lacking *mir* or *stau*^{11,15}. Thus, two important questions remain unanswered: how is Insc anchored in the apical cytocortex of the neuroblast; and how does the neuroblast acquire its apical–basal polarity?

We have shown that embryos mutant for the gene *bazooka* (*baz*) have defects in spindle orientation in neuroblasts, albeit with low penetrance³. *baz* encodes a cytoplasmic protein with three PDZ domains which has overall sequence similarity to Par-3 of *C. elegans* and rat ASIP^{3,17,18}. In neuroblasts, the spatial and temporal expression pattern of Baz is very similar to that of Insc^{1,2} (Fig. 1). During delamination, both proteins are colocalized in the stalk at the apical tip of the neuroblast, which makes contacts with adjacent epithelial cells (Fig. 1a, b). In pro- and metaphase, Baz and Insc are both

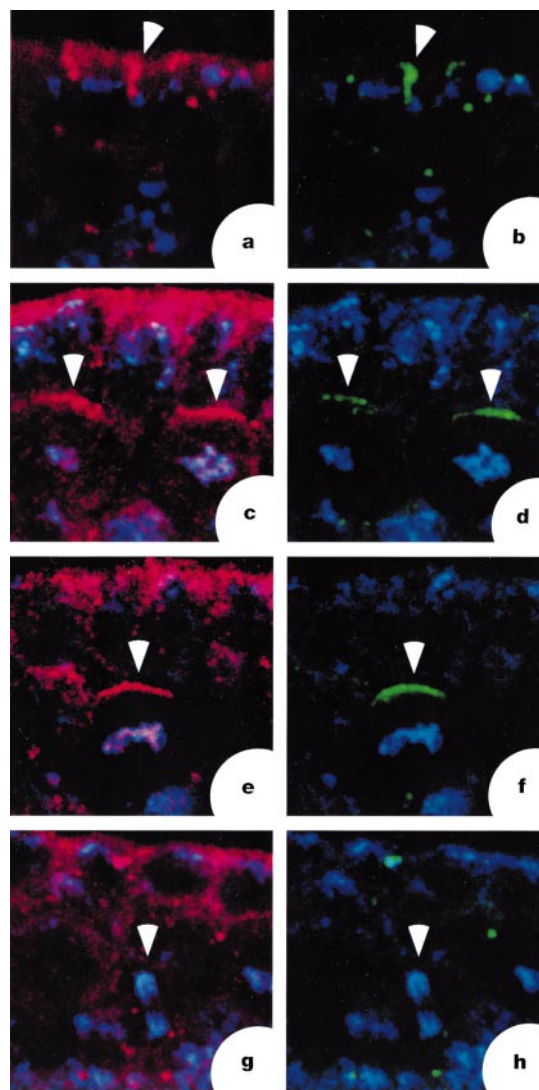


Figure 1 Baz and Insc colocalize in neuroblasts. Embryos at the extended germ-band stage were triply stained for Baz (red), Insc (green) and DNA (YOYO-1, blue). Baz/DNA staining is shown on the left (a, c, e, g) and Insc/DNA staining of the same neuroblasts on the right (b, d, f, h). During delamination (a, b), Baz and Insc colocalize in the stalk at the apical tip of the neuroblast (arrowhead) that is wedged between adjacent epithelial cells. In prophase (c, d) and metaphase (e, f), Baz and Insc form compact apical crescents which disappear during anaphase (g, h). Arrowhead marks the apical pole of the neuroblast. Apical is up in all panels.

localized in a crescent in the apical cortex, and become undetectable during anaphase (Fig. 1c–h). These findings indicated that the two proteins may interact directly with each other.

In *baz* mutant embryos lacking only the zygotic gene product, mislocalization of Insc occurred in only a small percentage of neuroblasts (data not shown). In contrast, embryos devoid of both maternal and zygotic Baz failed to form apical crescents of Insc altogether (Fig. 2b, compare Fig. 2a). Instead, Insc staining appeared diffuse and was not concentrated in the cell cortex (Fig. 2b). Embryos lacking maternal and zygotic Baz also showed a fully penetrant spindle orientation defect, very similar to that seen in *insc* mutants (data not shown). In neuroblasts of *insc* mutant embryos, Baz staining was greatly reduced (Fig. 2e, compare Fig. 2c); however, weak apical crescents of Baz were still detectable (Figure 2e, arrows, inset), indicating that localization of Baz may be independent of Insc. We found that the intensity of Baz staining in neuroblasts was closely correlated with the intensity of Insc staining; embryos

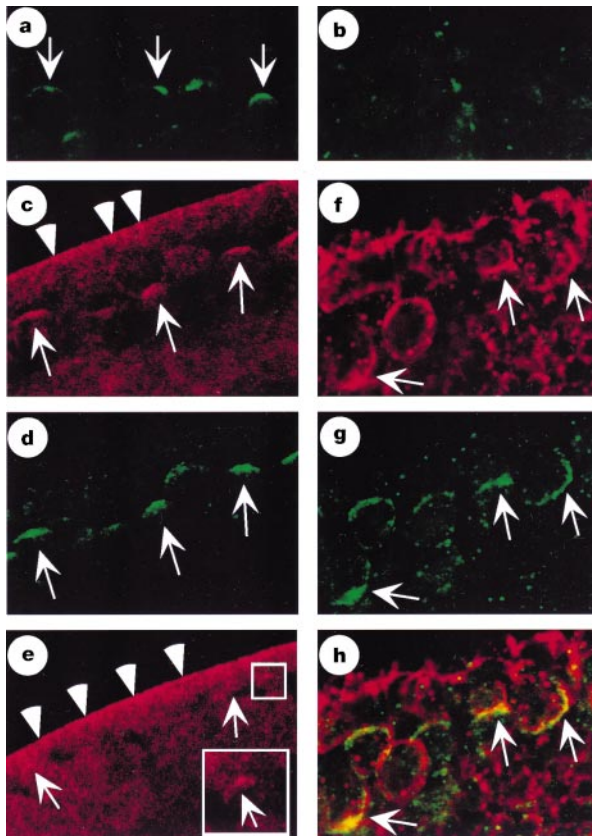


Figure 2 Baz is necessary and sufficient for cortical localization of Insc. **a, b**, Cortical localization of Insc is lost in *baz* mutants. **a**, Insc localization in a wild-type embryo at the extended germ-band stage. Arrows mark apical Insc crescents. **b**, In *baz*^{x106} mutant embryos derived from germ-line clones, Insc staining is diffuse and apical crescents do not form. **c–e**, Baz staining is reduced in neuroblasts of *insc* mutants. **c**, In heterozygous *insc*^{P49}/Cyo blue embryos, Baz is detectable in the apical region of the neuroectodermal epithelium (arrowheads) and in apical crescents in neuroblasts (arrows). **d**, Insc staining of the embryo shown in **c** reveals colocalization with Baz in neuroblasts. **e**, In homozygous *insc*^{P49} embryos, Baz staining in neuroblasts is reduced, whereas staining in the neuroectodermal epithelium is unaffected (arrowheads). Weak apical crescents are still detectable in some neuroblasts (inset, arrows). Inset, higher magnification of the boxed region in **e**. Homozygous mutant embryos were identified by the absence of both Insc and β -Gal staining. **f–h**, Overexpression of Baz causes mislocalization of Insc. **f**, Baz (red) is not restricted to the apical cortex when overexpressed. **g**, Concomitantly, Insc (green) colocalizes with Baz in ectopic positions (**h**, arrows). Note that some neuroblasts show strong Baz staining but little or no Insc staining. This is because Baz was expressed from a strong, constitutively active promoter that is not cell-cycle dependent, whereas the endogenous Insc is apparently degraded during anaphase. Apical is up in all panels.

heterozygous for an *insc* null mutation showed weaker Baz staining than wild type-embryos (data not shown). These findings may mean that Baz needs to associate with Insc to escape proteolytic degradation in neuroblasts.

To test whether Baz is sufficient for localization of Insc in neuroblasts, we overexpressed Baz ubiquitously using the GAL4 system. Ectopic localization of Baz occurred at random positions all around the cytocortex of neuroblasts (Fig. 2f). Concomitantly, Insc was targeted to the same ectopic positions (Fig. 2g, h), consistent with the idea that Baz is not only required but also sufficient for localization of Insc.

As one function of Insc is to localize Pros, and Insc localization was lost in *baz* mutant embryos, we checked whether Pros was also affected. In wild-type neuroblasts, Pros becomes localized to the apical cell cortex in late interphase and moves basally during mitosis^{5–7}. In metaphase, Pros forms a tight basal crescent. Upon completion of mitosis, Pros is found exclusively in the cortex of the budding GMC, and enters the nucleus of the GMC shortly after division^{5–7}. In *baz* mutant embryos derived from germ-line clones, Pros was found all around the cell cortex from late interphase to metaphase (Fig. 3a–f). We occasionally observed Pros crescents in metaphase, but their positions were random, as in *insc* mutants¹ (data not shown). Despite the lack of basal crescents in metaphase, Pros was enriched in the cortex of the budding GMC in late anaphase (Fig. 3g, h), leading to preferential segregation of Pros into the GMC. However, the position of the budding GMC was random with respect to the apical–basal axis of the neuroblast (Fig. 3a). Thus, Baz is required for the early steps of Pros localization which lead to the formation of basal crescents during metaphase. In contrast, the segregation of Pros into the budding GMC during anaphase and telophase appears to be controlled by a Baz-independent mechanism.

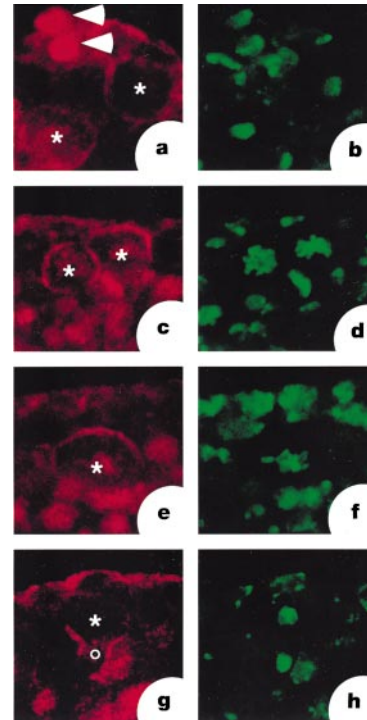


Figure 3 The subcellular localization of Pros is abnormal in *baz* mutant embryos. Pros staining is red (**a, c, e, g**) and DNA green (**b, d, f, h**). From late interphase (**a, b**) through prophase (**c, d**) and metaphase (**e, f**), Pros is found all around the cell cortex in most neuroblasts. During anaphase (**g, h**), Pros is preferentially localized in the cortex of the budding GMC. Stars, neuroblasts; circle, the budding GMC (**g**). Note that GMCs are often found in inappropriate positions close to the apical surface of the embryo (arrowheads in **a**). Apical is up in all panels.

Further insight into the relationship between Baz and Insc was obtained by analysing the localization of both proteins in *Drosophila* Schneider S2 cells. When expressed on its own, Insc was distributed diffusely in the cytoplasm (Fig. 4a), whereas Baz was concentrated in scattered clusters in the cytocortex (Fig. 4e). Co-expression of both proteins in the same cell resulted in redistribution of Insc to the cortex, where it colocalized with Baz (Fig. 4g–i). Strikingly, the localization of Baz in S2 cells was cell-cycle dependent, regardless of whether or not Insc was co-expressed. During interphase, Baz was found in discontinuous cortical patches (Fig. 4k, see also e, h), which were randomly scattered around the periphery of the cell. In contrast, we found a smoothly distributed cortical staining pattern in mitotic cells. In the latter, Baz often formed contiguous crescents (Fig. 4m), reminiscent of the staining pattern in mitotic neuroblasts (compare Fig. 1c and e). In doubly transfected cells, the localization of Insc followed exactly the cell-cycle-dependent changes in Baz localization (Fig. 4j, l, see also i). These data show that in S2 cells Baz always localizes to the cytocortex, whether Insc is present or not. In contrast, Insc does not become associated with the cytocortex unless Baz is expressed in the same cell, supporting the conclusion that Baz mediates cortical localization of Insc. Moreover, Baz forms cortical crescents in a cell-cycle-dependent manner and recruits Insc into these crescents. Thus, in S2 cells and in embryos, Baz is not only required but also sufficient for the asymmetric cortical localization of Insc.

Does the colocalization of Baz and Insc in neuroblasts and in S2 cells reflect a physical association between the two proteins? To test this hypothesis, cell extracts from wild-type *Drosophila* embryos, untransfected S2 cells, S2 cells transfected with either Insc or Baz and of doubly transfected cells were used for co-immunoprecipitation experiments. Using an antibody directed against Baz, a complex containing Insc was precipitated from embryos and from doubly transfected cells (Fig. 5a). A much weaker signal was detected in immunoprecipitates from cells transfected with Insc only, owing to low-level expression of endogenous Baz in S2 cells (Fig. 5a). In the reverse experiment, anti-Insc antibody precipitated a complex containing Baz from doubly transfected cells (Fig. 5b). These data show that Insc and Baz form a complex in wild-type embryos and when co-expressed in S2 cells.

Insc binds directly to Mir and Stau and is required for apical localization of these two proteins in neuroblasts^{10,11,15}. If Baz does anchor Insc in the apical cortex, then Mir and Stau should be found in the complex together with Baz and Insc. S2 cells express endogenous Staufen, but not Miranda (data not shown) and thus Miranda was not considered further in our analysis. Using an antibody against Stau for immunoprecipitation, a complex containing Baz was recovered from cells co-expressing Insc and Baz, but not from cells expressing Baz alone (Fig. 5c). The amount of coprecipitated Baz was much lower than in the co-immunoprecipitation with Insc (Fig. 5b), owing to the fact that the level of endogenous Stau in S2 cells is much lower than that of overexpressed Insc and Baz. Co-immunoprecipitation experiments using Baz antibody for immunoprecipitation and Stau antibody for western blot and binding experiments with a Baz–glutathione *S*-transferase (GST) fusion protein confirmed the association of Stau with Insc and Baz (data not shown). Thus, we conclude that Stau associated with Baz indirectly through Insc, and that Insc can form a complex with Baz and Stau at the same time. Together, our data indicate that Baz anchors a multiprotein complex containing Insc and Stau at the apical pole of neuroblasts.

Asymmetric cell division has been proposed to depend on an asymmetrically positioned cue in the mother cell¹⁹. This cue is necessary for the unequal distribution of other proteins, some of which are required to orient the mitotic spindle and hence guarantee the differential distribution of cytoplasmic factors to the daughter cells. Baz is a likely candidate for the postulated asymmetric cue. Baz is apically localized in epithelial cells and is necessary for the maintenance of epithelial polarity^{3,4}. We propose that Baz

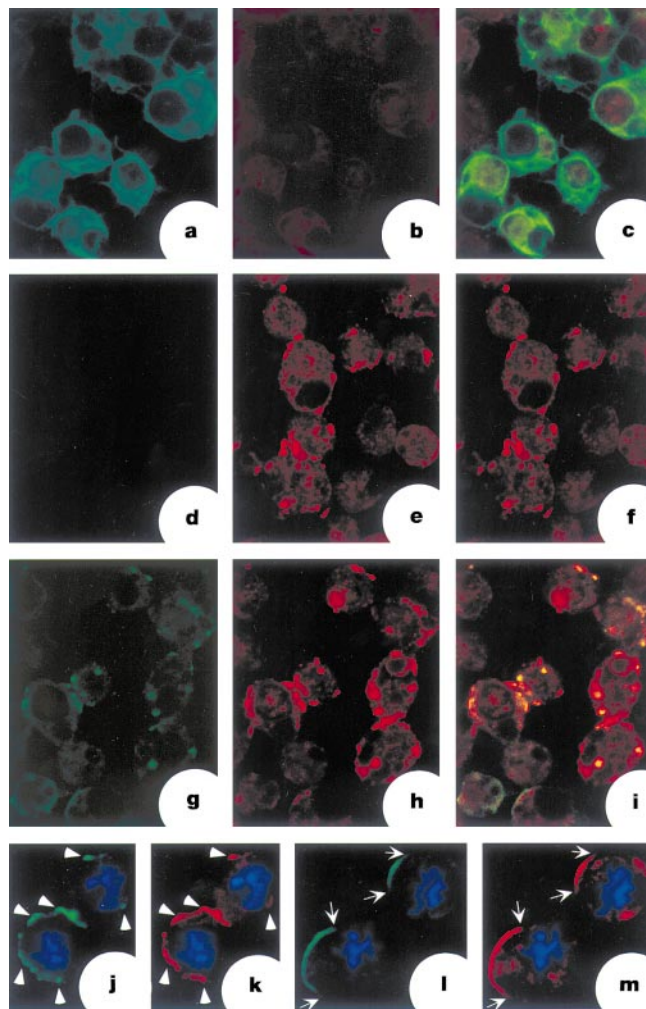


Figure 4 Cortical localization of Insc depends on Baz in S2 cells. **a–c**, S2 cells transfected with Insc, and stained for Insc (**a**) and Baz (**b**); **c**, merged image. Overexpressed Insc shows diffuse distribution in the cytoplasm. Endogenous Baz is barely detectable. **d–f**, S2 cells transfected with Baz, and stained for Insc (**d**) and Baz (**e**); **f**, merged image. Overexpressed Baz is in scattered cortical patches and Insc is not expressed. **g–i**, S2 cells doubly transfected with Insc and Baz, and stained for Insc (**g**) and Baz (**h**); **i**, merged image. Insc colocalizes with Baz in cortical patches, the diffuse cytoplasmic staining of Insc is reduced (compare **a** and **g**). **j–m**, Cell-cycle dependence of Baz and Insc localization in co-transfected S2 cells. In interphase (**j**, **k**), Insc (**j**, green) and Baz (**k**, red) are localized in scattered cortical patches (arrowheads). During mitosis (**l**, **m**), Insc and Baz form contiguous cortical crescents. Arrows mark crescent borders. DNA was stained with YOYO-1 (blue).

also maintains apico–basal positional information after delamination of neuroblasts by targeting Insc to the apical cytocortex. Insc, in turn, is required to reorient the axis of the mitotic spindle so that it is perpendicular to the spindle orientation in the overlying epithelial cells. As expected on the basis of this model, Insc is not found in epithelial cells of the ventral neurogenic region, and when ectopically expressed there, it localizes apically (possibly targeted by Baz) and induces a 90° rotation of the mitotic spindle¹.

The role of Baz during neuroblast division shows intriguing similarities to the role of Par-3 in early embryogenesis in *C. elegans*. Par-3 is localized asymmetrically to the anterior cortex of the zygote, where it is responsible for the asymmetric segregation of germ-line-specific P granules to the posterior P₁ cell and for correct positioning of the mitotic spindle¹⁷. Asymmetric localization of Par-3 requires another PDZ domain protein, Par-6, and a member of the protein kinase C family, PKC-3. All three proteins colocalize and are mutually dependent on each other for proper localization^{20,21}.

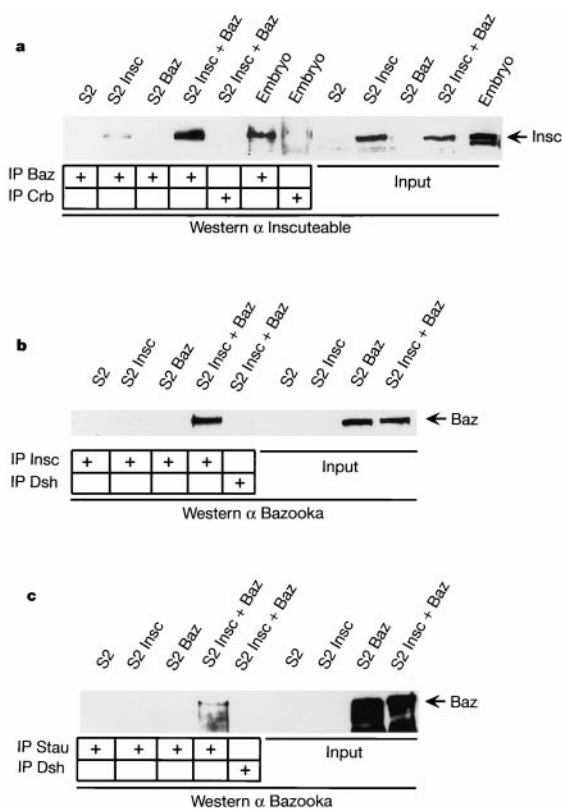


Figure 5 Baz, Insc and Stau are associated in a complex. **a**, Anti-Baz antibody (C-terminal, mouse) was used for immunoprecipitation (IP) of the embryonic and cell extracts indicated at the top. Immunoprecipitates underwent SDS-PAGE and western analysis with anti-Insc antibody (left lanes). An antibody against Crb²⁷ was used as a negative control. The western blot of the extracts used for immunoprecipitations (input) is shown on the right. **b**, In the converse experiment, anti-Insc antibody was used for immunoprecipitation and anti-Baz antibody (C-terminal, rat) for probing the western blot. An antibody against Dsh²⁵ was used as a negative control. **c**, Anti-Stau antibody was used for immunoprecipitations and anti-Baz antibody (N-terminal, rat) for western analysis. Compared with results in **b**, the ratio of the IP signal to the input signal is lower, most probably because the level of endogenous Stau is considerably lower than that of overexpressed Insc and Baz. Additional bands of lower molecular mass are specific breakdown products of Baz that are detected by the N-terminal anti-Baz antibody but not by the C-terminal anti-Baz antibody (see **b**).

Elimination of Par-6 or PKC-3 by mutation or by RNA interference causes phenotypes very similar to those observed in *par-3* mutants. Homologues of Par-6 and PKC-3 have been identified in *Drosophila*²⁰ (A.W. and E.K., unpublished data) and it will be exciting to see whether they also function in neuroblast division. □

Methods

Fly stocks and genetics

Oregon R was used as wild-type stock. *baz*^{X106} germ-line clone embryos were produced using the FLP/DFS technique⁴²². For analysis of Baz localization in *insc* mutants, *insc*^{P49} and *insc*^{P72} (ref. 2) were used in combination with a CyO blue balancer. Baz was overexpressed using the UAS-GAL4 system²³, using a UAS Baz transgene³ and the maternal GAL4 driver mat67G4 (D. St Johnston and J. P. Vincent, unpublished data), which drives early, ubiquitous expression in embryos.

Immunofluorescence staining of embryos and cells

Embryos were fixed in 3.7% formaldehyde, 100 mM phosphate buffer, pH 7.4. S2 cells were fixed in 3.7% formaldehyde, PBS. The primary antibodies used were rat anti-Baz N-term (1:200), rabbit anti-Insc² (1:1000), mouse anti-Pros MR1A⁷ (1:5). Images were taken on a Leica TCS NT confocal microscope and processed using Adobe Photoshop software.

Generation of Baz and Insc expression constructs and generation of antibodies

For expression of Baz and Insc in S2 cells, the respective full-length cDNAs were cloned into the pMK33 expression vector²⁴. For generation of antibodies against Baz, amino-

terminal (amino acids 1–297) and carboxy-terminal (amino acids 1328–1464 or 947–1464) portions of Baz were cloned into pGEX expression vectors (Pharmacia). The resulting Baz-GST fusion proteins were expressed in bacteria and used for immunization of mice and rats. Further details of the expression constructs will be provided on request.

Maintenance of S2 cells, preparation of cell lysates, immunoblotting and immunoprecipitation

Methods were as described²⁵. For western analysis, the antibodies used were rat anti-Baz C-term (1:1000), rat anti-Baz N-term (1:1000), rabbit anti-Insc² (1:4000). For immunoprecipitations, 2 µl of mouse anti-Baz C-term, 1 µl of rabbit anti-Insc or 2 µl of rabbit anti-Stau²⁶ were added to the cell or embryo lysates. Antibodies against Crb²⁷ or Dsh²⁵ were used as negative controls.

Received 12 August; accepted 1 October 1999.

- Kraut, R., Chia, W., Jan, L. Y., Jan, Y. N. & Knoblich, J. A. Role of inscuteable in orienting asymmetric cell divisions in *Drosophila*. *Nature* **383**, 50–55 (1996).
- Kraut, R. & Campos-Ortega, J. A. Inscuteable, a neural precursor gene of *Drosophila*, encodes a candidate for a cytoskeleton adaptor protein. *Dev. Biol.* **174**, 65–81 (1996).
- Kuchinke, U., Grawe, F. & Knust, E. Control of spindle orientation in *Drosophila* by the Par-3-related PDZ-domain protein Bazooka. *Curr. Biol.* **8**, 1357–1365 (1998).
- Müller, H. A. & Wieschaus, E. armadillo, bazooka, and Stardust are critical for early stages in formation of the zonula adherens and maintenance of the polarized blastoderm epithelium in *Drosophila*. *J. Cell. Biol.* **134**, 149–163 (1996).
- Knoblich, J. A., Jan, L. Y. & Jan, Y. N. Asymmetric segregation of Numb and Prospero during cell division. *Nature* **377**, 624–627 (1995).
- Hirata, J., Nakagoshi, H., Nabeshima, Y. & Matsuzaki, F. Asymmetric segregation of the homeo-domain protein Prospero during *Drosophila* development. *Nature* **377**, 627–630 (1995).
- Spana, E. P. & Doe, C. Q. The prospero transcription factor is asymmetrically localized to the cell cortex during neuroblast mitosis in *Drosophila*. *Development* **121**, 3187–3195 (1995).
- Knoblich, J. A. Mechanisms of asymmetric cell division during animal development. *Curr. Opin. Cell Biol.* **9**, 833–841 (1997).
- Jan, Y. N. & Jan, L. Y. Asymmetric cell division. *Nature* **392**, 775–778 (1998).
- Shen, C. P. et al. Miranda as a multidomain adapter linking apically localized Inscuteable and basally localized Staufer and Prospero during asymmetric cell division in *Drosophila*. *Genes Dev.* **12**, 1837–1846 (1998).
- Li, P., Yang, X., Wasser, M., Cai, Y. & Chia, W. Inscuteable and Staufer mediate asymmetric localization and segregation of prospero RNA during *Drosophila* neuroblast cell divisions. *Cell* **90**, 437–447 (1997).
- Ikeshima-Kataoka, H., Skeath, J. B., Nabeshima, Y., Doe, C. Q. & Matsuzaki, F. Miranda directs Prospero to a daughter cell during *Drosophila* asymmetric divisions. *Nature* **390**, 625–629 (1997).
- Matsuzaki, F., Ohshiro, T., Ikeshima-Kataoka, H. & Izumi, H. Miranda localizes staufer and prospero asymmetrically in mitotic neuroblasts and epithelial cells in early *Drosophila* embryogenesis. *Development* **125**, 4089–4098 (1998).
- Schuldt, A. J. et al. Miranda mediates asymmetric protein and RNA localization in the developing nervous system. *Genes Dev.* **12**, 1847–1857 (1998).
- Shen, C. P., Jan, L. Y. & Jan, Y. N. Miranda is required for the asymmetric localization of Prospero during mitosis in *Drosophila*. *Cell* **90**, 449–458 (1997).
- Broadus, J., Fuerstenberg, S. & Doe, C. Q. Staufer-dependent localization of prospero mRNA contributes to neuroblast daughter-cell fate. *Nature* **391**, 792–795 (1998).
- Etemad-Moghadam, B., Guo, S. & Kemphues, K. J. Asymmetrically distributed PAR-3 protein contributes to cell polarity and spindle alignment in early *C. elegans* embryos. *Cell* **83**, 743–752 (1995).
- Izumi, Y. et al. An atypical PKC directly associates and colocalizes at the epithelial tight junction with ASIP, a mammalian homologue of *Caenorhabditis elegans* polarity protein PAR-3. *J. Cell Biol.* **143**, 95–106 (1998).
- Gönczy, P. & Hyman, A. A. Cortical domains and the mechanism of asymmetric cell division. *Trends Cell Biol.* **6**, 382–387 (1996).
- Hung, T. J. & Kemphues, K. J. PAR-6 is a conserved PDZ domain-containing protein that colocalizes with PAR-3 in *Caenorhabditis elegans* embryos. *Development* **126**, 127–135 (1999).
- Tabuse, Y. et al. Atypical protein kinase C cooperates with PAR-3 to establish embryonic polarity in *Caenorhabditis elegans*. *Development* **125**, 3607–3614 (1998).
- Chou, T. B. & Perrimon, N. Use of a yeast site-specific recombinase to produce female germline chimeras in *Drosophila*. *Genetics* **131**, 643–653 (1992).
- Brand, A. H. & Perrimon, N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401–415 (1993).
- Koelle, M. R. et al. The *Drosophila* EcR gene encodes an ecdysone receptor, a new member of the steroid receptor superfamily. *Cell* **67**, 59–77 (1991).
- Willert, K., Brink, M., Wodarz, A., Varmus, H. & Nusse, R. Casein kinase 2 associates with and phosphorylates dishevelled. *EMBO J.* **16**, 3089–3096 (1997).
- St Johnston, D., Beuchle, D. & Nüsslein-Volhard, C. Staufer, a gene required to localize maternal RNAs in the *Drosophila* egg. *Cell* **66**, 51–63 (1991).
- Tepass, U., Theres, C. & Knust, E. crumbs encodes an EGF-like protein expressed on apical membranes of *Drosophila* epithelial cells and required for organization of epithelia. *Cell* **61**, 787–799 (1990).

Acknowledgements

We thank B. Chia, R. Kraut, D. St Johnston and C. Doe for antibodies; A. Müller and E. Wieschaus for *baz*^{X106} FRT 18D flies; D. St Johnston for mat67G4 flies; and J. Campos-Ortega for Insc cDNA. F. Grawe, M. Müller-Borg and A. Grimm provided technical assistance. We also thank J. Campos-Ortega, A. Müller, O. Bossinger, K. Johnson and P. Hardy for critical reading of the manuscript, and J. Knoblich for communication of results before publication. This work was supported by grants from the Deutsche Forschungsgemeinschaft.

Correspondence and requests for material should be addressed to A.W. (e-mail: wodarz@uni-duesseldorf.de).

Bazooka recruits Inscuteable to orient asymmetric cell divisions in *Drosophila* neuroblasts

Markus Schober*, Matthias Schaefer* & Juergen A. Knoblich

Research Institute of Molecular Pathology (I.M.P.), Dr Bohr Gasse 7, A-1030 Vienna, Austria

* These authors contributed equally to this work

Asymmetric cell divisions can be generated by the segregation of determinants into one of the two daughter cells^{1–4}. In *Drosophila*, neuroblasts divide asymmetrically along the apical–basal axis shortly after their delamination from the neuroectodermal epithelium. Several proteins, including Numb and Miranda, segregate into the basal daughter cell and are needed for the determination of its correct cell fate^{5–10}. Both the apical–basal orientation of the mitotic spindle and the localization of Numb and Miranda to the basal cell cortex are directed by Inscuteable, a protein that localizes to the apical cell cortex before and during neuroblast mitosis^{11,12}. Here we show that the apical localization of Inscuteable requires Bazooka, a protein containing a PDZ domain that is essential for apical–basal polarity in epithelial cells^{13,14}. Bazooka localizes with Inscuteable in neuroblasts and binds to the Inscuteable localization domain *in vitro* and *in vivo*. In embryos lacking both maternal and zygotic *bazooka* function, Inscuteable no longer localizes asymmetrically in neuroblasts and is instead uniformly distributed in the cytoplasm. Mitotic spindles in neuroblasts are misoriented in these embryos, and the proteins Numb and Miranda fail to localize asymmetrically in metaphase. Our results suggest that direct binding to Bazooka mediates the asymmetric localization of Inscuteable and connects the asymmetric division of neuroblasts to the axis of epithelial apical–basal polarity.

Drosophila neuroblasts delaminate from a polarized epithelium in the ventral neuroectoderm and divide asymmetrically along the apical–basal axis shortly after delamination. The protein Inscuteable is not expressed in the neuroectodermal epithelium, but expression starts in neuroblasts during delamination¹¹. The protein localizes into a crescent along the apical cell cortex until anaphase, when it becomes undetectable^{11,12}. In the absence of Inscuteable, mitotic spindles in neuroblasts are randomly oriented. The proteins Numb and Miranda still localize asymmetrically in these mutants but their crescents form at random positions and no longer overlie one of the two spindle poles^{9,12}. Expression and apical localization in metaphase neuroblasts have also been reported for Bazooka¹⁴. But unlike Inscuteable, Bazooka is expressed and apically localized in epithelial cells¹⁴. To compare Inscuteable and Bazooka localization directly through the cell cycle in neuroblasts, we stained stage 10 *Drosophila* embryos for Bazooka (Fig. 1a, c, e, g), Inscuteable (Fig. 1b, d, f, h) and DNA (Fig. 1a–h). During neuroblast delamination, both Inscuteable and Bazooka were detected in a stalk that is left behind in the epithelial cell layer (Fig. 1a, b). In delaminated neuroblasts, both proteins colocalized at the apical cell cortex during interphase (Fig. 1c, d) and metaphase of the first mitosis (Fig. 1e, f), but became undetectable during anaphase (Fig. 1g, h).

To test whether Bazooka and Inscuteable physically interact with each other, we performed co-immunoprecipitation experiments using a fly strain that ubiquitously expresses the Inscuteable central asymmetry domain fused to β -galactosidase (β -Gal) (Insc-cen- β -Gal) (Fig. 1k). This fusion protein can localize asymmetrically in neuroblasts and rescues the defects in spindle orientation and localization of Numb and Miranda in *inscuteable* mutants¹⁵. Flies expressing a non-functional part of Inscuteable fused to β -Gal (Insc-04- β -Gal)¹⁵ were used as control. Both Insc-cen- β -Gal and

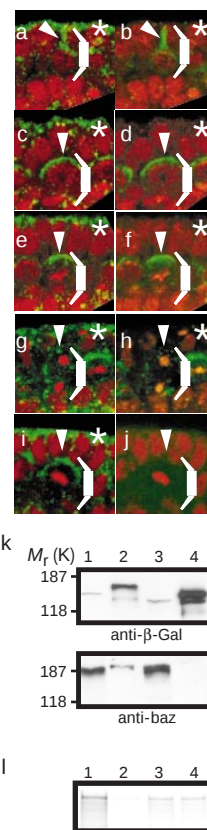


Figure 1 Localization and binding of Inscuteable and Bazooka. **a–j**, Stage-10 wild-type (**a–h**) and *inscuteable*^{P72} (**i, j**) embryos¹¹ were stained with rabbit anti-Inscuteable, rat anti-Bazooka and propidium iodide (for DNA). Apical is up and basal is down in this and all other figures. Brackets indicate the positions of neuroblasts. **a, c, e, g, i**, Bazooka (green) and DNA (red); **b, d, f, h, j**, Inscuteable (green) and DNA (red) in the same cells. Epidermal cells express Bazooka, but not Inscuteable (asterisks). In neuroblasts, both proteins (arrowheads) colocalize during delamination (**a, b**), in interphase (**c, d**) and in metaphase (**e, f**), but disappear during anaphase (**g, h**). In *inscuteable*^{P72} embryos no Inscuteable protein can be detected (**j**), but Bazooka is still apically localized (**i**). Note that the image in **i** was taken at a higher gain. The intensity of anti-Bazooka staining in neuroblasts (arrowheads) was strongly decreased in *inscuteable*^{P72} embryos, whereas it was unchanged in epidermal cells (asterisk). **k**, Protein extracts were prepared from embryos expressing Insc-cen- β -Gal¹⁵ (lanes 1 and 2) or Insc-04- β -Gal (lanes 3 and 4) under the control of the *hsp70* promoter. Immunoprecipitations were performed with anti- β -Gal (see Methods). Equal amounts of the precipitates were analysed by immunoblotting with anti- β -Gal (lanes 2 and 4) and anti-Bazooka (anti-baz, lanes 2 and 4). One-eighth of the amount of extract used for immunoprecipitation was loaded in lanes 1 and 3. A significant amount of Bazooka is precipitated by Insc-cen- β -Gal (lane 2) but not Insc-04- β -Gal (lane 4). **l**, ³⁵S-labelled full-length Bazooka protein translated *in vitro* was precipitated with GST (lane 2), GST–Inscuteable (lane 3) or GST–Insc-cen¹⁵ (lane 4) bound to glutathione beads. Precipitated proteins were separated by SDS–PAGE. One-tenth of the input was loaded in lane 1. Bazooka is precipitated by GST–Insc (lane 3) and GST–Insc-cen (lane 4) but not by GST alone (lane 2).

Insc-04- β -Gal could be immunoprecipitated with an anti- β -Gal antibody (Fig. 1k, anti- β -Gal, lanes 2 and 4, respectively). Whereas a significant amount of Bazooka protein was immunoprecipitated together with Insc-cen- β -Gal (Fig. 1k, anti-baz, lane 2), no Bazooka protein could be detected in the Insc-04- β -Gal immunoprecipitate (Fig. 1k, anti-baz, lane 4), suggesting that Bazooka specifically binds either directly or indirectly to the Inscuteable asymmetry domain *in vivo*. To test a possible direct interaction between Bazooka and Inscuteable, full-length Bazooka protein translated *in vitro* was incubated with beads containing glutathione S-transferase (GST) alone, GST fused to full-length Inscuteable or GST fused to the

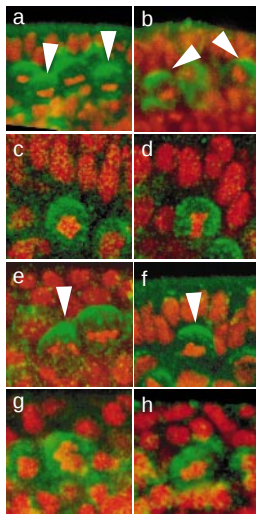


Figure 2 Defects in Inscuteable localization in *bazooka* embryos. Stage-10 embryos were stained with rabbit anti-Inscuteable (green) and propidium iodide (for DNA, red). **a, b**, Inscuteable crescents form apically in wild-type neuroblasts (arrowheads in **a**) but are frequently mislocalized in *bazooka*^{B15-8} embryos²⁴ (arrowheads in **b**). **c, d**, *bazooka*^{W106} mutant germline clones were produced as described¹³ with the Flp-DFS system²⁵. No Inscuteable crescents were detected in prophase (**c**) or metaphase (**d**) neuroblasts from embryos lacking maternal and zygotic *bazooka* function (*bazooka*^{GLC} embryos). **e–h**, RNAi experiments. In control injected embryos (**e, f**), Inscuteable localizes asymmetrically to the apical neuroblast cortex in prophase (**e**) and metaphase (**f**) (arrowheads). In *bazooka*^{RNAi} embryos (**g, h**), Inscuteable is not asymmetrically localized in neuroblasts and is found mostly in the cytoplasm.

Inscuteable asymmetry domain (Fig. 1l). Bazooka could bind to full-length Inscuteable (lane 3) and the Inscuteable asymmetry domain (lane 4), but not to GST alone (lane 2). Taken together, these results suggest that Bazooka and Inscuteable directly bind to each other *in vivo*.

The binding and colocalization of Inscuteable and Bazooka indicate a functional connection between the two proteins. We therefore stained *inscuteable*-null mutant embryos for Bazooka, Inscuteable and DNA (Fig. 1i, j). Even though no Inscuteable protein could be detected in these embryos (Fig. 1j), Bazooka was still localized to the apical cell cortex (Fig. 1i). Anti-Bazooka staining was much weaker in neuroblasts but remained unchanged in epidermal cells of these embryos (note that Fig. 1i is taken at a higher gain than Fig. 1e). This indicates that Inscuteable is required for the stabilization of Bazooka in neuroblasts, whereas another, unknown and Inscuteable-independent, mechanism stabilizes

Bazooka in epithelial cells. We also analysed the subcellular distribution of Inscuteable in *bazooka* mutant embryos (Fig. 2a, b). Even though Inscuteable still localized into an apical crescent in all observed neuroblasts in these mutants (Fig. 2b), this crescent did not form at the apical cortex in some neuroblasts but was instead found at a more lateral position (arrowheads in Fig. 2b). The fact that Inscuteable is still localized in *bazooka* mutants could be due to maternally contributed RNA.

Indeed, when Inscuteable localization was analysed in *bazooka* mutant germline clones, no apical crescents were observed in neuroblasts from embryos lacking both maternal and zygotic *bazooka* function (*bazooka*^{GLC} embryos) (Fig. 2c, d), and most of the Inscuteable protein was uniformly distributed in the cytoplasm instead. Only a small number of embryos were recovered from *bazooka* germline clones. To exclude the possibility that the observed phenotypes were actually secondary consequences of defects occurring during oogenesis, maternal and zygotic *bazooka* RNA were specifically disrupted in embryos by using double-stranded-RNA (dsRNA)-mediated interference (RNAi)^{16–19}. No Bazooka protein could be detected by immunofluorescence in stage-10 embryos injected with *bazooka* dsRNA (*bazooka*^{RNAi} embryos; see Methods), whereas protein levels were normal in buffer-injected control embryos (data not shown). In control embryos, Inscuteable was correctly localized to the apical neuroblast cell cortex (Fig. 2e, f) and apical–basal polarity in epithelial cells was unaffected (data not shown). In *bazooka*^{RNAi} embryos, epithelial cells lost their apical–basal polarity (data not shown) and Inscuteable was not asymmetrically localized in neuroblasts but was evenly distributed in the cytoplasm instead (Fig. 2g, h). We conclude from these experiments that *bazooka* is essential for asymmetric and cortical localization of Inscuteable. In contrast, *inscuteable* is dispensable for Bazooka localization even though it may be required for Bazooka protein stability.

To analyse the requirement of Bazooka for Miranda and Numb localization in neuroblasts, control embryos (Fig. 3a, b, g, h) and *bazooka*^{RNAi} embryos (Fig. 3c, d, i, j) were stained for Miranda (Fig. 3a–d) or Numb (Fig. 3g–j) and for DNA. As in *inscuteable* mutant embryos, metaphase plates were frequently misoriented in *bazooka*^{RNAi} embryos (Fig. 3c). Both Miranda (Fig. 3c) and Numb (Fig. 3i) were distributed homogeneously around the cell cortex in all metaphase neuroblasts from *bazooka*^{RNAi} embryos, whereas all metaphase neuroblasts from control embryos showed asymmetric localization of Miranda (Fig. 3a) and Numb (Fig. 3g). Similar results were obtained in *bazooka*^{GLC} embryos (Fig. 3e, k). Despite these defects in Miranda and Numb localization during metaphase, both proteins were often partially localized over one of the two spindle poles in anaphase neuroblasts (Fig. 3d, f, j, l) and often but not always found at a higher concentration in one of the two daughter

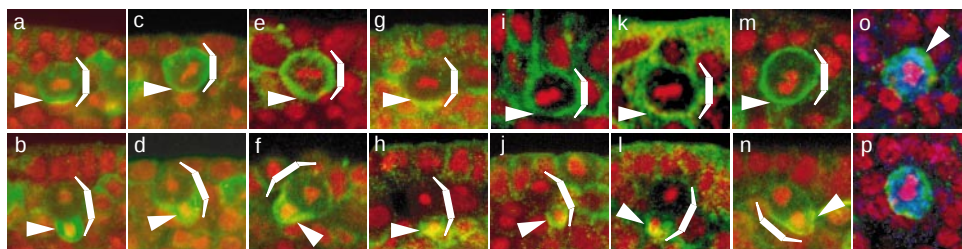


Figure 3 Defects in Miranda and Numb localization in the absence of Bazooka and Inscuteable. **a–n**, Control (**a, b, g, h**), *bazooka*^{RNAi} (**c, d, i, j**), *bazooka*^{GLC} (**e, f, k, l**) and *inscuteable*^{P72} (**m, n**) embryos were stained with rabbit anti-Miranda (**a–f, m, n**; green) or rabbit anti-Numb (**g–l**; green), and propidium iodide (red). Arrowheads indicate the basal neuroblast cell cortex. Miranda and Numb are asymmetrically localized in control metaphase neuroblasts (**a, g**) and segregate into the basal daughter cell in telophase (**b, h**). In *bazooka*^{RNAi} (**c, d, i, j**), *bazooka*^{GLC} (**e, f, k, l**) and *inscuteable*^{P72} (**m, n**)

neuroblasts, both proteins are evenly distributed at the cell cortex in metaphase (**c, e, i, k, m**) but concentrate over the basal spindle pole in telophase (**d, f, j, l, n**). **o, p**, Embryos were stained with rabbit anti-Miranda (green), DAPI (for DNA, red) and guinea-pig anti-Asense (blue) to identify SOP cells. Miranda is asymmetrically localized in metaphase SOP cells from control embryos (**o**) but not from *bazooka*^{RNAi} (**p**) embryos.

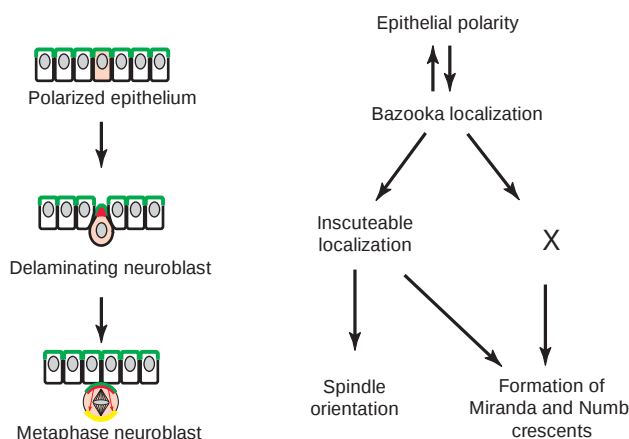


Figure 4 Model for Bazooka and Inscuteable function during asymmetric cell division in neuroblasts. Apically localized Bazooka protein is inherited by delaminating neuroblasts. Inscuteable localizes apically by binding to Bazooka and directs apical–basal neuroblast

division. Comparison of the *inscuteable* and *bazooka* phenotypes suggests a second, Inscuteable-independent, function for Bazooka (X).

cells after mitosis (data not shown). Asymmetric distribution of Numb and Miranda during late mitosis in *inscuteable* mutants has not been described before, but a careful re-examination of Numb and Miranda localization in anaphase and telophase revealed a similar asymmetric distribution in both *inscuteable* mutants (data not shown) and *inscuteable* mutant embryos injected with *bazooka* dsRNA (creating *bazooka inscuteable* double mutants; Fig. 3m, n). We conclude that Bazooka is essential for the asymmetric localization of Numb and Miranda during metaphase, whereas a second mechanism acts in anaphase and telophase to concentrate Numb and Miranda in the cortical area overlying one of the two spindle poles even in the absence of Inscuteable and Bazooka.

Numb and Miranda are also asymmetrically localized in sensory organ precursor (SOP) cells^{5,9}, the progenitors of the peripheral nervous system. Even though Inscuteable is expressed in the peripheral nervous system¹¹, it is not required for Numb and Miranda localization in SOP cells (ref. 12; and data not shown). Bazooka was also expressed and asymmetrically localized in SOP cells (data not shown). To test Bazooka function in these cells, we stained control and *bazooka*^{RNAi} embryos for Miranda, DNA and Asense. SOP cells were identified on the basis of position and *asense* expression. In all Miranda expressing prophase and metaphase SOP cells the protein was asymmetrically localized in late stage-10 control embryos (Fig. 3o), whereas no asymmetrical localization was detected in *bazooka*^{RNAi} embryos (Fig. 3p). These results suggest that Bazooka functions in SOP cells via an Inscuteable-independent mechanism.

The mechanisms that orient asymmetric cell division are well understood in yeast²⁰. A G-protein cascade interprets a mark that is left behind at the cortex after cytokinesis and initiates bud formation next to this site, thus orienting the axis of division relatively to the previous cell division²¹. A similar cell division remnant has been suggested to be involved in orienting asymmetric divisions in *Caenorhabditis elegans*²². Our results suggest that *Drosophila* neuroblasts use a different mechanism to coordinate asymmetric cell divisions with the general body plan (Fig. 4). Bazooka protein is apically localized in epithelial cells, where it is required for maintaining apical–basal polarity. The apical localization of Bazooka is maintained when some epithelial cells assume a neuronal fate and delaminate to become neuroblasts. Neuroblasts start expressing Inscuteable, which binds to Bazooka, localizes apically and directs the apical–basal orientation of asymmetric cell division. Thus, Bazooka and Inscuteable cooperate in connecting asymmetric cell division to the axis of epithelial polarity.

Although Numb and Miranda still form crescents (even though at

random positions) in many *inscuteable* mutant neuroblasts, no asymmetric localization during metaphase can be detected in the absence of Bazooka. Non-localized Inscuteable protein could cause more severe defects than the absence of the protein. But it is more likely that Bazooka also has an Inscuteable-independent function either directly in the localization of Miranda and Numb or of another molecule that is functionally related to Inscuteable (X in Fig. 4). Consistent with this second role is the observed Bazooka function in SOP cells, which do not require Inscuteable. However, unlike neuroblasts, SOP cells do not divide with an apical–basal orientation. It is possible that they reorient their polarity during delamination. Alternatively, Bazooka might have a function in these cells that is not related to apical–basal polarity. Bazooka encodes the *Drosophila* homologue of the *C. elegans* gene *par-3* (ref. 14), which is required for establishing polarity during the first cell divisions of the zygote²³. Analogously to Bazooka, Par-3 localizes asymmetrically during these divisions and is required for spindle orientation and the asymmetric segregation of P-granules into one of the two daughter cells²³. Surprisingly, database searches of the sequenced *C. elegans* genome do not reveal any sequence homologue of Inscuteable, suggesting that other genes function downstream of Par-3 to direct asymmetric cell divisions. It is intriguing to speculate that this second pathway also exists downstream of Bazooka and accounts for the differences between the *inscuteable* and *bazooka*^{RNAi} phenotypes in neuroblasts that we observe. □

Methods

Immunohistology

Immunofluorescence experiments were performed as described⁵ using rabbit anti-Inscuteable¹² (1 : 1000 dilution), mouse anti-β-Gal (Promega, 1 : 100), rabbit anti-Miranda⁹ (1 : 1000), guinea-pig anti-Asense⁵ (1 : 5000) and rat anti-Bazooka (1 : 300, a gift from A. Wodarz) as primary antibodies, and Alexa-488 (Molecular Probes), Cy3 or Cy5 (Jackson) conjugated secondary antibodies. Rabbit anti-Numb was generated against a peptide that has been described⁵, affinity-purified and used at a dilution of 1 : 100. Images were recorded on a Leica TCS-NT confocal microscope equipped with a Coherent-Vitesse pulsed near-infrared laser to visualize DAPI (4',6-diamidino-2-phenylindole dihydrochloride) by two-photon-excitation microscopy.

RNAi experiments

Nucleotides 3623–4664 and nt 1944–3248 of the *bazooka* transcript (numbers from the published cDNA sequence¹⁴) were PCR-amplified from *Drosophila* embryonic first-strand cDNA by using primers that introduced T7 promoters at both ends. DNA fragments were gel-purified and transcriptions *in vitro* were performed with the Megascript kit (Ambion). The quality and quantity of the RNA were controlled on agarose gels. RNA injections were performed as described¹⁷ except that devitellinization with methanol was used instead of devitellinization by hand. Both *bazooka* fragments yielded identical results.

Protein binding assays

Embryos (2–5 hours old) from the fly stocks *hs-Gal4, UAS-insc-cen-β-Gal/TM3* (expressing residues 252–615 fused to β-Gal¹⁵) and *hs-Gal4, UAS-insc-04-β-Gal/TM3* (expressing residues 436–552 fused to β-Gal¹⁵) were heat-shocked for 30 min at 37 °C. After recovery for 1 h at 25 °C, protein extracts were prepared in extraction buffer (25 mM Tris, pH 8.0, 27.5 mM NaCl, 20 mM KCl, 25 mM sucrose, 10 mM EDTA, 10 mM EGTA, 1 mM dithiothreitol, 10% (v/v) glycerol, 0.5% Nonidet P40) containing protease and phosphatase inhibitors and centrifuged for 15 min at 20,000 g. The supernatant was used in immunoprecipitations with anti-β-Gal coupled to Protein-A–Sepharose (Pharmacia). Beads were washed three times (10 min each) in extraction buffer. Bound proteins were analysed by immunoblotting with mouse anti-β-Gal (Promega) and rat anti-Bazooka. For binding *in vitro*, a 4.5 kb DNA fragment containing the complete *bazooka* open reading frame was PCR-amplified from *Drosophila* embryonic first-strand cDNA, cloned into pBluescript and sequenced. Translations *in vitro* were performed with the TNT kit (Promega) and *in vitro* binding assays to GST–Inscuteable and GST–Insc-cen *in vitro* were done as described¹⁵.

Received 13 August; accepted 26 October 1999.

- Horvitz, H. R. & Herskowitz, I. Mechanisms of asymmetric cell division: two Bs or not two Bs, that is the question. *Cell* **68**, 237–255 (1992).
- Lu, B., Jan, L. Y. & Jan, Y. N. Asymmetric cell division: lessons from flies and worms. *Curr. Opin. Genet. Dev.* **8**, 392–399 (1998).
- Jan, Y. N. & Jan, L. Y. Asymmetric cell division. *Nature* **392**, 775–778 (1998).
- Knoblich, J. A. Mechanisms of asymmetric cell division during animal development. *Curr. Opin. Cell Biol.* **9**, 833–841 (1997).
- Rhyu, M. S., Jan, Y. N. Asymmetric distribution of Numb protein during division of the sensory organ precursor cell confers distinct fates to daughter cells. *Cell* **76**, 477–491 (1994).
- Spana, E. P., Kopczyński, C., Goodman, C. S. & Doe, C. Q. Asymmetric localization of Numb autonomously determines sibling neuron identity in the *Drosophila* CNS. *Development* **121**, 3489–3494 (1995).
- Spana, E. P. & Doe, C. Q. The Prospero transcription factor is asymmetrically localized to the cell cortex during neuroblast mitosis in *Drosophila*. *Development* **121**, 3187–3195 (1995).
- Knoblich, J. A., Jan, L. Y. & Jan, Y. N. Asymmetric segregation of Numb and Prospero during cell division. *Nature* **377**, 624–627 (1995).
- Shen, C. P., Jan, L. Y. & Jan, Y. N. Miranda is required for the asymmetric localization of Prospero during mitosis in *Drosophila*. *Cell* **90**, 449–458 (1997).
- Matsuzaki, F., Ohshiro, T., Ikeshima-Kataoka, H. & Izumi, H. Miranda localizes Stauf and Prospero asymmetrically in mitotic neuroblasts and epithelial cells in early *Drosophila* embryogenesis. *Development* **125**, 4089–4098 (1998).
- Kraut, R. & Campos-Ortega, J. A. *inscuteable*, a neural precursor gene of *Drosophila*, encodes a candidate for a cytoskeleton adaptor protein. *Dev. Biol.* **174**, 65–81 (1996).
- Kraut, R., Chia, W., Jan, L. Y., Jan, Y. N. & Knoblich, J. A. Role of *inscuteable* in orienting asymmetric cell divisions in *Drosophila*. *Nature* **383**, 50–55 (1996).
- Muller, H. A. & Wieschaus, E. *armadillo*, *bazooka* and *stardust* are critical for early stages in formation of the zonula adherens and maintenance of the polarized blastoderm epithelium in *Drosophila*. *J. Cell Biol.* **134**, 149–163 (1996).
- Kuchinke, U., Grawe, F. & Knust, E. Control of spindle orientation in *Drosophila* by the Par-3-related PDZ-domain protein Bazooka. *Curr. Biol.* **8**, 1357–1365 (1998).
- Knoblich, J. A., Jan, L. Y. & Jan, Y. N. Deletion analysis of the *Drosophila* *inscuteable* protein reveals domains for cortical localization and asymmetric localization. *Curr. Biol.* **9**, 155–158 (1999).
- Fire, A. *et al.* Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806–811 (1998).
- Kennerdell, J. R. & Carthew, R. W. Use of dsRNA-mediated genetic interference to demonstrate that *frizzled* and *frizzled 2* act in the wingless pathway. *Cell* **95**, 1017–1026 (1998).
- Misquitta, L. & Paterson, B. M. Targeted disruption of gene function in *Drosophila* by RNA interference (RNA-i): a role for *nautilus* in embryonic somatic muscle formation. *Proc. Natl Acad. Sci. USA* **96**, 1451–1456 (1999).
- Bhat, M. A. *et al.* Discs Lost, a novel multi-PDZ domain protein, establishes and maintains epithelial polarity. *Cell* **96**, 833–845 (1999).
- Chant, J. Generation of cell polarity in yeast. *Curr. Opin. Cell Biol.* **8**, 557–565 (1996).
- Madden, K. & Snyder, M. Cell polarity and morphogenesis in budding yeast. *Annu. Rev. Microbiol.* **52**, 687–744 (1998).
- Waddle, J. A., Cooper, J. A. & Waterston, R. H. Transient localized accumulation of actin in *Caenorhabditis elegans* blastomeres with oriented asymmetric divisions. *Development* **120**, 2317–2328 (1994).
- Etemad-Moghadam, B., Guo, S. & Kemphues, K. J. Asymmetrically distributed PAR-3 protein contributes to cell polarity and spindle alignment in early *C. elegans* embryos. *Cell* **83**, 743–752 (1995).
- McKim, K. A., Dahmus, J. B. & Hawley, R. S. Cloning of the *Drosophila melanogaster* meiotic recombination gene *mei-218*: a genetic and molecular analysis of interval 15E. *Genetics* **144**, 215–228 (1996).
- Chou, T. B. & Perrimon, N. Use of a yeast site-specific recombinase to produce female germline chimeras in *Drosophila*. *Genetics* **131**, 643–653 (1992).

Acknowledgements

We thank M. Glotzer for stimulating discussions; G. Christofori, B. Dickson and M. Glotzer for helpful comments on the manuscript; E. Kleiner for technical assistance; P. Steinlein and K. Paiha for help with confocal microscopy; W. Chia and Y. N. Jan for providing antibodies; K. S. McKim and A. Müller for providing fly stocks; and A. Wodarz and E. Knust for communicating results before publication.

Correspondence and requests for materials should be addressed to J.A.K. (e-mail: knoblich@nt.imp.univie.ac.at).

A telomerase component is defective in the human disease dyskeratosis congenita

James R. Mitchell, Emily Wood & Kathleen Collins

Division of Biochemistry and Molecular Biology, Department of Molecular and Cell Biology, 401 Barker Hall, University of California, Berkeley, California 94720-3204, USA

The X-linked form of the human disease dyskeratosis congenita (DKC) is caused by mutations in the gene encoding dyskerin¹. Sufferers have defects in highly regenerative tissues such as skin and bone marrow, chromosome instability and a predisposition to develop certain types of malignancy. Dyskerin is a putative pseudouridine synthase, and it has been suggested that DKC may be caused by a defect in ribosomal RNA processing. Here we show that dyskerin is associated not only with H/ACA small nucleolar RNAs², but also with human telomerase RNA, which contains an H/ACA RNA motif³. Telomerase adds simple sequence repeats to chromosome ends using an internal region of its RNA as a template⁴, and is required for the indefinite proliferation of primary human cells⁵. We find that primary fibroblasts and lymphoblasts from DKC-affected males are not detectably deficient in conventional H/ACA small nucleolar RNA accumulation or function; however, DKC cells have a lower level of telomerase RNA, produce lower levels of telomerase activity and have shorter telomeres than matched normal cells. The pathology of DKC is consistent with compromised telomerase function leading to a defect in telomere maintenance, which may limit the proliferative capacity of human somatic cells in epithelia and blood.

Mammalian telomerase RNAs differ from their ciliate counterparts by the presence of a 3' domain resembling an H/ACA small nucleolar RNA (snoRNA), an RNA family required for pseudouridine modification and precursor processing of rRNA³. Because the known mammalian telomerase protein components were identified on the basis of homology with their ciliate counterparts⁶, proteins that may interact with the human telomerase RNA (hTR) H/ACA domain have yet to be identified. In *Saccharomyces cerevisiae*, H/ACA snoRNAs are associated with several proteins, including the pseudouridine synthase Cbf5p (refs 7, 8). Mammalian orthologues of Cbf5p include the rat protein NAP57 (ref. 9) and human dyskerin, the product of the X-linked DKC locus¹.

To investigate whether dyskerin is a component of both conventional H/ACA small nucleolar ribonucleoproteins (snoRNPs) and the human telomerase ribonucleoprotein (RNP), we expressed dyskerin with an HA-epitope tag or human telomerase reverse transcriptase component (hTERT) with a Flag-epitope tag by transient transfection of telomerase-positive human 293 cells. Immunopurification specifically and efficiently recovered HA-dyskerin on HA-antibody resin, or Flag-hTERT on Flag-antibody resin (data not shown). Various RNAs were assayed for co-immunopurification by northern blot hybridization (Fig. 1a). HA-antibody immunopurification of HA-dyskerin-containing extract (lanes 3, 4) but not Flag-hTERT-containing extract (lanes 1, 2) recovered co-transfected recombinant hTR or co-transfected recombinant U64 H/ACA snoRNA. Co-transfected *Tetrahymena thermophila* telomerase RNA, which lacks an H/ACA domain, was not recovered on either antibody resin (data not shown). Endogenous H/ACA snoRNAs including hTR, U64, E1 and E2 (Fig. 1a) and U17 (data not shown) also co-purified specifically with HA-dyskerin. Endogenous RNAs were recovered by immunopurification less efficiently than recombinant RNAs, presumably due to transfection of only 25–50% of the cell population and the stability of pre-assembled

RNPs. As a control for immunopurification specificity, we checked for other RNAs, including the predominantly cytoplasmic signal recognition particle (SRP) RNA, the small nuclear RNA (snRNA) U2 and the box C/D snoRNA U3, all of which did not co-purify with HA-dyskerin on HA-antibody resin (Fig. 1a). As an additional control, Flag-antibody immunopurification of Flag-hTERT extract (lanes 5, 6) but not HA-dyskerin extract (lanes 7, 8) recovered both recombinant and endogenous hTERT but not any other RNA examined.

We assayed the immunopurified samples for telomerase activity to determine whether dyskerin remained part of a fully assembled,

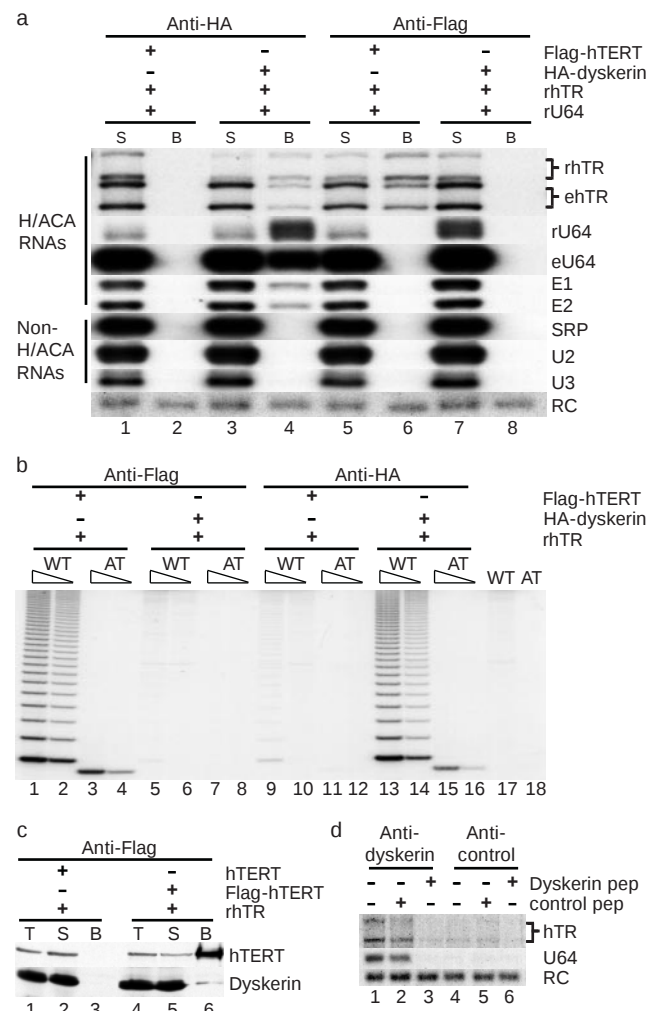


Figure 1 Co-purification of telomerase RNA, telomerase activity, hTERT and dyskerin. **a**, Supernatant (S) and HA- or Flag-antibody-bound (B) fractions from the indicated transfection extracts were analysed by northern blot hybridization. B samples were 2× relative to S samples. Recombinant (r) hTERT and U64 migrated with slower mobility than their endogenous (e) counterparts because of their sequence tags. hTERT migrates as a doublet, apparently because of partial folding³. A recovery control (RC) was added to samples before RNA extraction. **b**, Telomerase activity was determined in the antibody-bound samples from **a**. PCR primers were specific for telomerase products of the endogenous hTERT wild-type template (WT) or the recombinant hTERT altered template (AT). Each sample was assayed at 1× and 0.2× concentration. Lanes 17 and 18 were PCR controls with no extract and WT or AT primers. **c**, Total extracts (T), immunopurification supernatants (S) or antibody-bound (B) fractions from cells expressing transfected recombinant hTERT and either untagged hTERT (lanes 1–3) or Flag-hTERT (lanes 4–6) were analysed by immunoblot for hTERT (top) and endogenous dyskerin (bottom). B samples were 12.5× relative to T and S samples. **d**, Northern blot hybridization was performed to detect endogenous hTERT and U64 in samples immunopurified on dyskerin antibody resin (lanes 1–3) or control antibody resin (lanes 4–6) in the presence of the indicated peptide competitors. A recovery control (RC) was added before RNA extraction.

active telomerase RNP (Fig. 1b). HA-antibody immunopurification of HA-dyskerin extract but not Flag-hTERT extract recovered strong telomerase activity templated by endogenous hTERT (WT; lanes 13–14, 9–10) or the altered template of the co-transfected recombinant hTERT used in this experiment (AT; lanes 15–16, 11–12). PCR amplification of telomerase products synthesized with the altered template generates a single product instead of the wild-type 6-base-pair (bp) ladder^{3,10}. This change in the banding pattern is due, at least in part, to differences in PCR conditions rather than telomerase product synthesis, and so the wild-type and altered-template telomerase levels cannot be compared directly. As a control, Flag-antibody immunopurification of Flag-hTERT extract but not HA-dyskerin extract also recovered both endogenous hTERT- and recombinant hTERT-templated telomerase activities (lanes 1–8).

To confirm the association of endogenous dyskerin with telomerase, we first checked for the co-purification of endogenous dyskerin with recombinant Flag-hTERT immunopurified from transfected cell extracts. We carried out Flag-antibody immunopurification in parallel of extracts containing either transfected untagged hTERT or Flag-tagged hTERT. Flag-hTERT, but not untagged hTERT, was depleted from immunopurification supernatants and recovered on Flag-antibody resin (Fig. 1c, top). Using an affinity-purified polyclonal antibody raised against a dyskerin peptide, we detected the purification of endogenous dyskerin with Flag-antibody resin only in the presence of Flag-hTERT (Fig. 1c, bottom). Finally, we confirmed the association of endogenous dyskerin with endogenous hTERT. Using the dyskerin antibody described above, endogenous U64 H/ACA snoRNA and hTERT were both immunopurified from HeLa cell extract (Fig. 1d, lane 1). The association of hTERT and U64 with the dyskerin-antibody resin was blocked by an excess of the dyskerin peptide against which the antibody was raised (lane 3), but not by an unrelated peptide (lane 2). The levels of RNA associated with the dyskerin antibody resin in the presence of competitor dyskerin peptide (lane 3) were similar to background binding to a nonspecific affinity-purified, polyclonal, anti-peptide antibody resin (lanes 4–6). Samples immunopurified on dyskerin-antibody resin also had telomerase activity (data not shown). We conclude that dyskerin is both a conventional H/ACA snoRNP component and a telomerase RNP component, and that telomerase RNPs containing dyskerin are catalytically active.

Genetic depletion of the *S. cerevisiae* dyskerin homologue Cbf5p destabilizes H/ACA snoRNAs and inhibits H/ACA snoRNA-directed pseudouridine formation and precursor rRNA processing⁷. This led to the hypothesis that the phenotypes of DKC are induced by a ribosome biogenesis deficiency¹¹, which we tested by examining all commercially available DKC cell lines for potential defects in pseudouridine formation and precursor rRNA processing. These cell lines included two distinct cell types: lymphoblasts transformed with Epstein–Barr virus (EBV) and primary skin fibroblasts. Both were obtained as a clonal, early passage cultures. Lymphoblasts were available from a carrier mother (Carrier 1; these cells expressed the wild-type dyskerin allele) and her three clinically affected sons (DKC 1a, 1b and 1c; these cells expressed a mutant dyskerin allele) in family 1, and also a clinically affected male in family 2 (DKC 2; these cells expressed a different mutant dyskerin allele, see Methods). Fibroblasts were available from the clinically affected male in family 2 (DKC 2), his mother (Carrier 2a) and his maternal grandmother (Carrier 2b). All the cell lines above were mortal and eventually senesced in culture, including the EBV-transformed lymphoblasts which were not immortalized by the virus. Immortal cell lines analysed as controls included 3714, an EBV-transformed and immortalized lymphoblastoid cell line; 293, the adenovirus-transformed human embryonic kidney cell line used in our transfections; and VA13, an SV40-transformed human lung fibroblast cell line which maintains telomeres by a telomerase-independent mechanism¹².

We assayed for the presence of pseudouridines in rRNA from

DKC, carrier and control cell lines (Fig. 2a). Using a chemical modification assay¹³ shown to be sensitive to H/ACA snoRNA depletion in yeast¹⁴, we determined pseudouridine levels at several sites in 18S and 28S rRNAs where pseudouridine formation is directed by known H/ACA snoRNAs¹⁵. In this assay, pseudouridine modification blocks reverse transcription, resulting in a strong stop one nucleotide short of the pseudouridine position. In RNA from all cells examined, pseudouridine formation occurred at the expected sites including those targeted by E3, U23, U65, U68 and U69 (Fig. 2a) and U19 (data not shown). We also confirmed the presence of pseudouridine at a number of other known modification sites without identified cognate H/ACA snoRNAs (Fig. 2a; and data not shown). Normalization of the pseudouridine-imposed reverse-transcription stops to the run-off product of 18S rRNA reverse transcription (Fig. 2a, top) or to nonspecific reverse-transcription stops (data not shown) demonstrated quantitatively that rRNAs from all cells contained similar amounts of each H/ACA snoRNP-specified pseudouridine examined. We then looked for an accumulation of partially or abnormally processed rRNA. Equal amounts of total RNA from representative DKC, carrier and control cell lines had similar amounts and ratios of 18S and 28S rRNAs (Fig. 2b). No partially processed intermediates were observed in any of several RNA preparations, suggesting that H/ACA snoRNP function in

rRNA processing is either not required in humans or not affected by DKC isoforms of dyskerin.

The lack of a discernible defect in rRNA modification or processing in DKC cells, together with the absence of evidence for a potentially redundant dyskerin protein family, suggests that DKC mutations do not substantially inhibit accumulation of dyskerin protein or conventional H/ACA snoRNAs. Indeed, our dyskerin antibody recognized a polypeptide of similar mobility and abundance in DKC, carrier and control cells (Fig. 2c). In addition, northern blot hybridization demonstrated that at least 4 out of the 13 identified human H/ACA snoRNAs were present at similar levels in total RNA from all cells examined (U64, E1 and E2, Fig. 3a; U19, data not shown). Slight variations in accumulation of these RNAs occurred but with no correlation to the allele of dyskerin expressed. Non-H/ACA RNAs including SRP RNA, U2 snRNA and U3 box C/D snoRNA were also present at comparable levels (Fig. 3a).

In contrast, the accumulation of hTR was reduced in all cells expressing a DKC isoform of dyskerin (Fig. 3a, top). The steady-state level of hTR can vary between cell types (compare 293 cells,

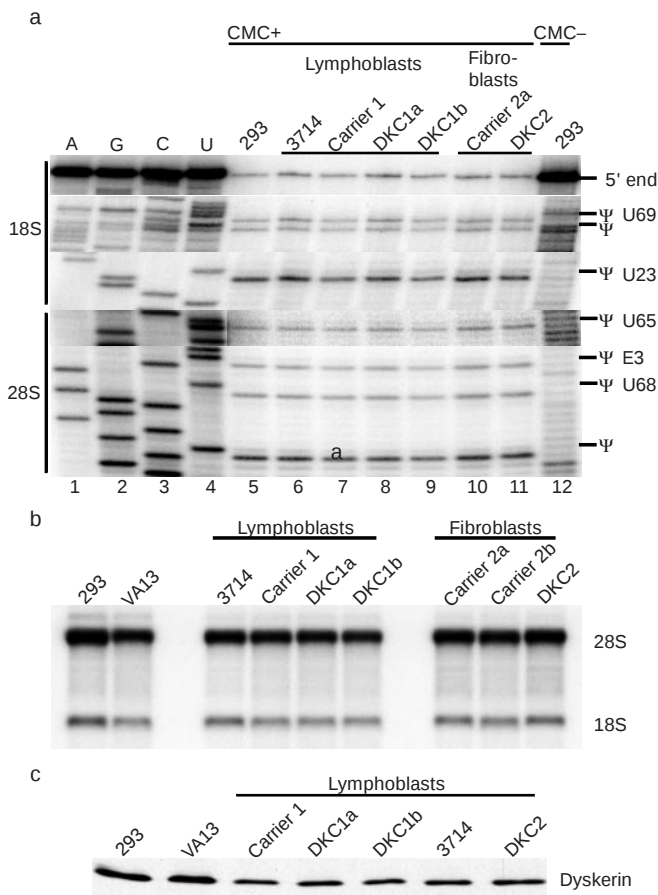


Figure 2 Analysis of rRNA and dyskerin. **a**, Pseudouridines were detected by reverse-transcriptase-mediated primer extension of equal amounts of unmodified RNA (lane 12) or modified RNA from various cell types (lanes 5–11). In lanes 1–4, reverse transcription of unmodified 293 cell RNA was done in the presence of a ddNTP to generate a ladder complementary to the rRNA sequence indicated at top. Previously mapped pseudouridines Ψ and H/ACA snoRNAs that target specific sites are indicated. **b**, 28S and 18S rRNAs were examined by northern blot hybridization using equal amounts of total RNA. **c**, Endogenous dyskerin in the indicated cell lysates was analysed by immunoblot. Loading was normalized by cell number and each lane had an equivalent amount of protein, as judged by staining with Ponceau S.

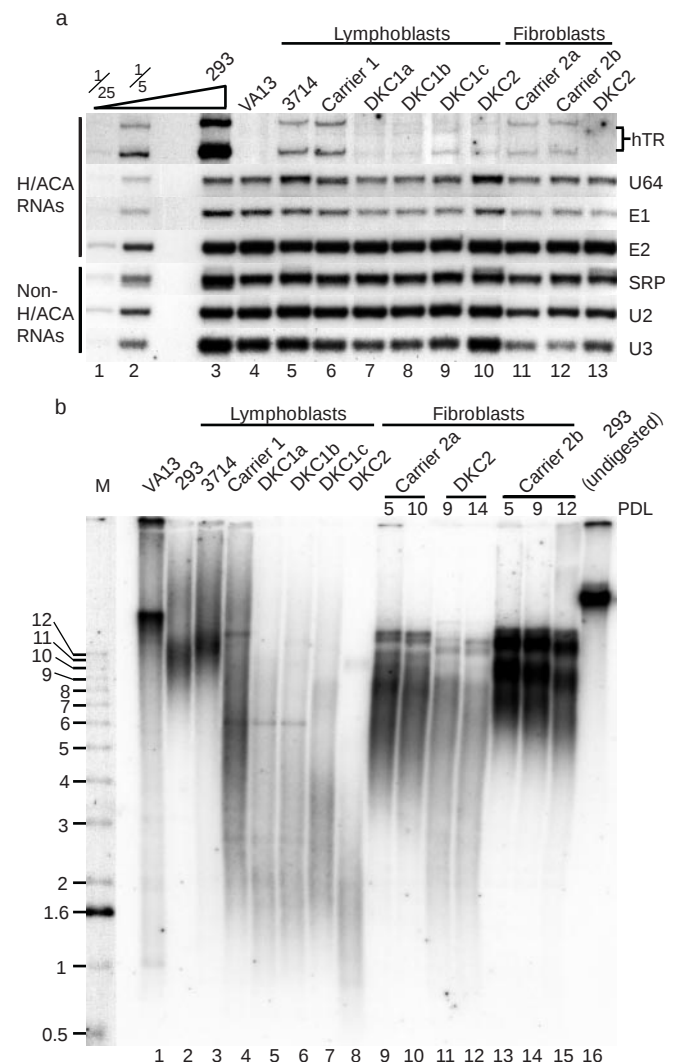


Figure 3 Reduced hTR and short telomeres in DKC cells. **a**, Levels of hTR and other RNAs were examined by northern blot hybridization. Lanes 3–13 contained 1 μ g of total RNA; lanes 1 and 2, 40 and 200 ng, respectively. **b**, Genomic DNA was restriction-digested, resolved by electrophoresis, and analysed by in-gel hybridization with the radiolabelled telomeric-repeat sequence oligonucleotide (TTAGGG)₃. Either 100 ng (lanes 1–3, 16) or 500 ng (lanes 4–15) of DNA was loaded. M, migration of molecular weight standards in kbp. PDL, population doubling at which DNA was prepared.

normal lymphoblasts and normal fibroblasts in lanes 1–3, 5–6 and 11–12, respectively). However, comparison of DKC cells to carrier or control cells of the same cell type revealed a roughly fivefold reduction in hTR. The lymphoblasts from the four affected males in two families (lanes 7–10) all had less hTR than the carrier (lane 6) and control (lane 5) lymphoblasts. Similarly, fibroblasts from the DKC-affected male in family 2 (lane 13) had less hTR than fibroblasts of either carrier (lanes 11, 12). This correlation between expression of a DKC dyskerin isoform and reduced hTR level was evident in multiple, independent preparations of RNA from both lymphoblasts and fibroblasts (data not shown).

We examined telomere length in DKC, carrier and control cells. Because all of the DKC and carrier lymphoblasts and fibroblasts lacked strong telomerase activity during their limited passaging in culture (data not shown), their telomere lengths should reflect the relative balance of proliferation-dependent telomere loss and telomerase activity over their proliferative history *in vivo*. Terminal restriction fragments (TRFs) in lymphoblasts from the three DKC-affected brothers at ages 15, 18 and 19 (Fig. 3b, lanes 5–7, respectively) were several kbp shorter than TRFs in lymphoblasts from their clinically unaffected, 47-year-old mother (lane 4). This finding is in striking contrast with the decline in TRF length versus donor age seen in a variety of cell types^{16,17}. TRFs in lymphoblasts from the unrelated DKC-affected male at age 7 were also abnormally short (lane 8). TRFs in all the DKC lymphoblasts were much shorter than expected for age-matched normal individuals¹⁶ and in fact were at or below the critically short telomere length that induces growth crisis in culture^{18,19}. Immortal cells maintained much longer telomeres by telomerase-dependent (lanes 2–3) or telomerase-independent (lane 1) mechanisms. In the fibroblasts from family 2, TRFs in cells from the 7-year-old DKC patient (lanes 11–12) were shorter than in cells from his mother (Carrier 2a, lanes 9–10) or his maternal grandmother (Carrier 2b, lanes 13–15). All fibroblast TRFs shortened with proliferation (lanes 9–15), confirming the telomeric origin of the hybridization signal. When the blot was stripped and re-probed for minisatellite DNA, hybridizing restriction fragments did not shorten with cell proliferation and confirmed the relatedness among the members of family 1 or family 2 (data not shown).

Most human primary cells accumulate hTR but have undetectable or low telomerase activity due to limiting expression of hTERT. Because fibroblasts that constitutively express hTERT from a strong promoter gain telomerase activity⁵, we used hTERT overexpression to investigate whether DKC fibroblasts produced less functional

telomerase than carrier fibroblasts. Equal numbers of DKC and carrier fibroblasts from family 2 were transfected with expression constructs for hTERT alone or hTERT and hTR. An immunoblot of transfected cell extracts (Fig. 4a, top) revealed roughly equal expression of hTERT among the cell lines transfected with hTERT alone (lanes 2, 5, 8) or with hTERT and hTR (lanes 3, 6, 9). The immunoblot was re-probed for endogenous dyskerin (Fig. 4a, bottom), which revealed a slight underloading of lane 6. With or without co-expression of hTR, which appeared to increase hTERT accumulation, hTERT levels were much higher than in even the most strongly telomerase-positive cell lines (data not shown).

We determined telomerase activity in a series of fivefold dilutions of each of the transfected cell extracts (Fig. 4b). Without transfected hTERT expression, none of the fibroblasts had detectable telomerase activity (dilution series 1, 4, and 7). With overexpression of hTERT alone, both carrier fibroblasts had about fivefold more telomerase activity than the DKC fibroblasts (compare dilution series 2, 5 and 8) in proportion to their relative levels of hTR (Fig. 3a). This difference was reproduced in four independent experiments with either an untagged or Flag-tagged hTERT expression construct (data not shown). Co-transfection of hTERT and wild-type template hTR expression vectors increased telomerase activity in all three fibroblast cultures (dilution series 3, 6 and 9). This increase was reproducibly most marked in the DKC fibroblasts, which attained a level of telomerase activity close to that of the carrier fibroblasts despite the expression of a DKC isoform of dyskerin.

We cannot, at present, rule out the possibility that a defect in ribosome biogenesis, or some other cellular process, contributes to the pathology of human DKC. However, the known phenotypic hallmarks of DKC include skin abnormalities and deficiencies of the haematopoietic system^{20,21}—defects in highly proliferative tissues that could ensue from a telomere maintenance disorder. Mortality occurs primarily from bone marrow failure associated with reduced regenerative capacity of haematopoietic stem cells²². As the disease progresses, unbalanced chromosomal rearrangements including dicentric and ring chromosomes, and Robertsonian translocations become detectable in skin, bone marrow and peripheral blood cells, despite the unaltered sensitivity of DKC cells to DNA damaging agents²³. How might expression of a mutant dyskerin gene inhibit telomerase RNP accumulation? An aberrant dyskerin molecule could affect RNP assembly on the hTR H/ACA domain, hTR 3' end processing, recruitment of other protein components onto hTR or RNP turnover. Because several features of the telomerase RNP are unique amongst H/ACA RNPs, a DKC isoform of dyskerin could compromise telomerase function without affecting conventional H/ACA snoRNPs. Further work will be required to clarify the mechanism of DKC dyskerin influence on telomerase. □

Methods

Constructs and cell lines

HA-dyskerin was expressed in pRcCMV (Invitrogen). Expression constructs for Flag-hTERT and untagged hTERT were gifts from Amgen and the Weinberg laboratory, respectively. Constructs for the expression of wild-type and altered-template hTR and for U64 have been described³. DKC cell lines were obtained from the Coriell Cell Repository. DKC family 1 has been clinically described^{24,25}, and the dyskerin messenger RNA expressed in affected individuals resulted in the single amino-acid substitution T66A (data not shown). The carrier cells from this family expressed wild-type dyskerin mRNA (data not shown). DKC family 2 carried a single amino-acid deletion of L37 (ref. 1). EBV-infected lymphoblasts were received without previous passaging. Fibroblasts from Carrier 2a (age 30), Carrier 2b (age 78) and DKC 2 (age 7) were received at 3, 8 and 7 population doublings respectively. DKC, carrier and 3714 lymphoblasts were grown in RPMI supplemented with 15% fetal bovine serum (FBS), sodium pyruvate and antibiotics. Other cells were grown in DMEM with 10% FBS and antibiotics.

Transfection, extract preparation, immunopurification and telomerase activity assay

Transient transfection of 293 cells was done as described³. Fibroblasts were transfected using Lipofectamine Plus (Gibco) with 1 µg of expression vectors or empty vectors combined per well of a 6-well plate. Transfection of control expression vectors was equally

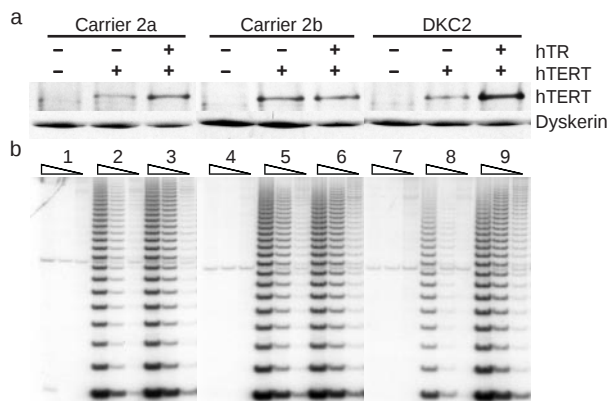


Figure 4 Reduced telomerase activity in DKC fibroblasts. Fibroblasts from two female carriers and one DKC-affected male from the same family were transfected with expression constructs for hTERT alone or with wild-type template hTR as indicated. **a**, Recombinant hTERT (top) and endogenous dyskerin (bottom) in cell lysates were detected by immunoblot. Lane 6 was slightly underloaded, as judged by Ponceau S staining. **b**, Telomerase activity in transfected cell extracts was assayed in three consecutive fivefold dilutions.

efficient for all fibroblasts. Extracts were made by freeze-thaw lysis as follows: cells were removed from plates with trypsin, washed in PBS, resuspended in hypotonic lysis buffer (20 mM HEPES pH 8.0, 2 mM MgCl₂, 0.2 mM EGTA, 10% glycerol, 1 mM DTT, 0.1 mM PMSF), and subjected to two consecutive freeze-thaw cycles from liquid nitrogen to a 37°C water bath. NaCl was added to 420 mM (293 cells) or 100 mM (fibroblasts) before clearing the lysate in a microcentrifuge for 15–20 min. HeLa cell extract was prepared by dounce homogenization in hypotonic buffer and cleared as above. HA and dyskerin antibodies were prebound to GammaBind G-Sepharose (Pharmacia) in wash buffer (hypotonic lysis buffer plus 100 mM NaCl and 0.1% NP-40) with blocking agents (100 ng µl⁻¹ BSA, 100 ng µl⁻¹ casein, 100 ng µl⁻¹ tRNA, 250 ng µl⁻¹ yeast total RNA, 100 ng µl⁻¹ glycogen). Flag antibody resin (Kodak) was blocked as above before use. Immunopurification was done at room temperature for 1 h in the presence of 0.1% NP-40. Peptide competitors were added to extract at 50 µg ml⁻¹ before addition of antibody resin. Telomerase activity assays were performed according to a modified³ TRAP²⁶ protocol.

DNA, RNA and protein analysis

Genomic DNA was digested overnight with *RsaI* and *HinfI*, resolved in 0.5% agarose/0.6 × TBE and analysed by in-gel hybridization²⁷. Probes for northern blot hybridization included 2'-O-methyl-substituted RNA oligonucleotides complementary to telomerase RNA (h5) (ref. 3) and DNA oligonucleotides complementary to all other RNAs. Pseudouridine analysis was done by *N*-cyclohexyl-*N'*-β-(4-methylmorpholinium)ethylcarbodiimide *p*-tosylate (CMC) modification and reverse transcription¹³. Reverse transcription reactions with three different primers were each performed at least two times. Antibody to dyskerin was raised and affinity purified against the peptide ERDTPYPRKWGLGPKASQ with a cysteine at the N terminus. Antibody to hTERT was raised and affinity purified against the peptide GSRHNERFLRNT with a cysteine at the N terminus.

Received 2 August; accepted 15 October 1999.

- Heiss, N. S. *et al.* X-linked dyskeratosis congenita is caused by mutations in a highly conserved gene with putative nucleolar functions. *Nature Genet.* **19**, 32–38 (1998).
- Tollervey, D. & Kiss, T. Function and synthesis of small nucleolar RNAs. *Curr. Opin. Cell Biol.* **9**, 337–342 (1997).
- Mitchell, J. R., Cheng, J. & Collins, K. A box H/ACA small nucleolar RNA-like domain at the human telomerase RNA 3' end. *Mol. Cell. Biol.* **19**, 567–576 (1999).
- Yu, G., Bradley, J. D., Attardi, L. D. & Blackburn, E. H. *In vivo* alteration of telomere sequences and senescence caused by mutated *Tetrahymena* telomerase RNAs. *Nature* **344**, 126–132 (1990).
- Colgin, L. M. & Reddel, R. R. Telomere maintenance mechanisms and cellular immortalization. *Curr. Opin. Genet. Dev.* **9**, 97–103 (1999).
- Nugent, C. I. & Lundblad, V. The telomerase reverse transcriptase: components and regulation. *Genes Dev.* **12**, 1073–1085 (1998).
- Lafontaine, D. L. J., Bousquet-Antonelli, C., Henry, Y., Caizergues-Ferrer, M. & Tollervey, D. The box H + ACA snoRNAs carry Cbf5p, the putative rRNA pseudouridine synthase. *Genes Dev.* **12**, 527–537 (1998).
- Watkins, N. J. *et al.* Cbf5p, a potential pseudouridine synthase, and Nhp2p, a putative RNA-binding protein, are present together with Gar1p in all H BOX/ACA-motif snoRNPs and constitute a common bipartite structure. *RNA* **4**, 1549–1568 (1998).
- Meier, U. T. & Blobel, G. NAP57, a mammalian nucleolar protein with a putative homolog in yeast and bacteria. *J. Cell Biol.* **127**, 1505–1514 (1994).
- Feng, J. *et al.* The RNA component of human telomerase. *Science* **269**, 1236–1241 (1995).
- Luzzatto, L. & Karadimitris, A. Dyskeratosis and ribosomal rebellion. *Nature Genet.* **19**, 6–7 (1998).
- Bryan, T. M., Englezou, A., Gupta, J., Bacchetti, S. & Reddel, R. R. Telomere elongation in immortal human cells without detectable telomerase activity. *EMBO J.* **14**, 4240–4248 (1995).
- Bakin, A. & Ofengand, J. Four newly located pseudouridylation residues in *Escherichia coli* 23S ribosomal RNA are all at the peptidyltransferase center: analysis by the application of a new sequencing technique. *Biochemistry* **32**, 9754–9762 (1993).
- Ni, J., Tien, A. & Fournier, M. Small nucleolar RNAs direct site-specific synthesis of pseudouridine in ribosomal RNA. *Cell* **89**, 565–573 (1997).
- Ganot, P., Bortolin, M.-L. & Kiss, T. Site-specific pseudouridine formation in preribosomal RNA is guided by small nucleolar RNAs. *Cell* **89**, 799–809 (1997).
- Hastie, N. D. *et al.* Telomere reduction in human colorectal carcinoma and with ageing. *Nature* **346**, 866–868 (1990).
- Harley, C. B., Futcher, A. B. & Greider, C. W. Telomeres shorten during ageing of human fibroblasts. *Nature* **345**, 458–460 (1990).
- Counter, C. M., Botelho, F. M., Wang, P., Harley, C. B. & Bacchetti, S. Stabilization of short telomeres and telomerase activity accompany immortalization of Epstein–Barr virus-transformed human B lymphocytes. *J. Virol.* **68**, 3410–3414 (1994).
- Tahara, H. *et al.* Abnormal telomere dynamics of B-lymphoblastoid cell strains from Werner's syndrome patients transformed by Epstein–Barr virus. *Oncogene* **15**, 1911–1920 (1997).
- Drachtman, R. A. & Alter, B. P. Dyskeratosis congenita: clinical features and genetic heterogeneity. *J. Pediatr. Hematol. Oncol.* **14**, 297–304 (1992).
- Dokai, I. Dyskeratosis congenita: an inherited bone marrow failure syndrome. *Br. J. Haematol.* **92**, 775–779 (1996).
- Marsh, J. C. W. *et al.* "Stem cell" origin of the hematopoietic defect in dyskeratosis congenita. *Blood* **79**, 3138–3144 (1992).
- Dokai, I. *et al.* Dyskeratosis congenita fibroblasts are abnormal and have unbalanced chromosome rearrangements. *Blood* **80**, 3090–3096 (1992).
- Sirinavin, C. & Trowbridge, A. A. Dyskeratosis congenita: clinical features and genetic aspects. *J. Med. Genet.* **12**, 339–354 (1975).
- Trowbridge, A. A., Sirinavin, C. & Linman, J. W. Dyskeratosis congenita: hematologic evaluation of a sibship and review of the literature. *Am. J. Hematol.* **3**, 143–152 (1977).
- Kim, N. W. *et al.* Specific association of human telomerase activity with immortal cells and cancer. *Science* **266**, 2011–2014 (1994).
- Prowse, K. R. & Greider, C. W. Developmental and tissue-specific regulation of mouse telomerase and telomere length. *Proc. Natl. Acad. Sci. USA* **92**, 4818–4822 (1995).

Acknowledgements

We thank A. Fischer for expert tissue culture assistance; Amgen and the Weinberg laboratory for their hTERT expression constructs; and D. Rio, M. Botchan, J. Rine and members of the Collins laboratory for comments on the manuscript. J.R.M. is a predoctoral fellow of the N.S.F.

Correspondence and requests for materials should be addressed to J.R.M. (e-mail: jmitch@uclink4.berkeley.edu) or K.C. (e-mail: kcollins@socrates.berkeley.edu).

The RCAF complex mediates chromatin assembly during DNA replication and repair

Jessica K. Tyler*, Christopher R. Adams†, Shaw-Ree Chen*‡, Ryuji Kobayashi§, Rohinton T. Kamakaka‡ & James T. Kadonaga*

*Department of Biology and Center for Molecular Genetics, University of California at San Diego, La Jolla, California 92093-0347, USA

†National Institutes of Health, NICHD, 18 Library Drive, Bethesda, Maryland 20892-5430, USA

§Cold Spring Harbor Laboratory, P.O. Box 100, Cold Spring Harbor, New York 11724, USA

Chromatin assembly is a fundamental biological process that is essential for the replication and maintenance of the eukaryotic genome^{1–4}. In dividing cells, newly synthesized DNA is rapidly assembled into chromatin by the deposition of a tetramer of the histone proteins H3 and H4, followed by the deposition of two dimers of histones H2A and H2B to complete the nucleosome—the fundamental repeating unit of chromatin⁵. Here we describe the identification, purification, cloning, and characterization of replication-coupling assembly factor (RCAF), a novel protein complex that facilitates the assembly of nucleosomes onto newly replicated DNA *in vitro*. RCAF comprises the *Drosophila* homologue of anti-silencing function 1 protein ASF1⁶ and histones H3 and H4. The specific acetylation pattern of H3 and H4 in RCAF is identical to that of newly synthesized histones. Genetic analyses in *Saccharomyces cerevisiae* demonstrate that ASF1 is essential for normal cell cycle progression, and suggest that RCAF mediates chromatin assembly after DNA replication and the repair of double-strand DNA damage *in vivo*.

The assembly of chromatin on to plasmids undergoing simian virus 40 (SV40) origin-dependent DNA replication can be recreated *in vitro* by using a crude extract that contains all of the factors required for replication except SV40 large T-antigen, and all of the factors required for chromatin assembly except chromatin assembly factor-1 (CAF-1)^{7–9}. To identify novel factors involved in the assembly of newly replicated DNA into chromatin, we have modified the conditions of the SV40 DNA-replication–chromatin-assembly system by supplementing the reactions with additional template DNA and exogenous CAF-1 (to ensure that CAF-1 is not a limiting component). Under these reaction conditions, there is no chromatin assembly (Fig. 1a, lane 1; chromatin assembly is measured by the introduction of negative supercoiling into plasmid DNA), presumably owing to the sequestration of an essential chromatin assembly factor or histones from the crude extract by the extra template DNA. The addition of factors involved in chromatin assembly (CAF-1, dCAF-1 p55¹⁰, ACF (ATP-utilizing chromatin assembly and remodelling factor)¹¹, NAP-1 (nucleosome assembly protein-1)¹², dNLP (*Drosophila* nucleoplasmin-like

‡Present address: Department of Biochemistry, University of Washington, Seattle, Washington 98195, USA

protein)¹³, or histones) failed to restore chromatin assembly activity (data not shown). However, the addition of a crude *Drosophila* embryo extract (S190 extract) resulted in complementation of the chromatin assembly activity in a replication-specific manner (Fig. 1a; replication-specific chromatin assembly is demonstrated by the preferential supercoiling of the newly replicated DNA relative to the bulk DNA). These results suggested that the S190 extract contains a novel chromatin assembly factor.

By using the modified reaction conditions as a complementation assay ('RCAF assay'), we purified this chromatin assembly activity from the S190 extract and obtained a novel factor that we have termed RCAF (Fig. 1b). Three polypeptides were correlated with the

peak of RCAF activity through the three final purification steps (Fig. 1c, d, fractions 5 and 6; and data not shown). We isolated the cDNA encoding the largest RCAF polypeptide by protein micro-sequencing. We have termed this protein dASF1 because it is encoded by a novel *Drosophila* gene that is homologous to the ASF1 (anti-silencing function 1) gene from *S. cerevisiae*. ASF1 was previously identified as a late G₁ phase-specific gene that derepresses transcriptional silencing when overexpressed from a plasmid^{6,14,15}. The amino-terminal 155 amino-acid residues of ASF1 share 67% identity between *Drosophila* and yeast, as well as 84% identity between *Drosophila* and an apparent ASF1 homologue in humans (constructed from two overlapping human expressed sequence tag

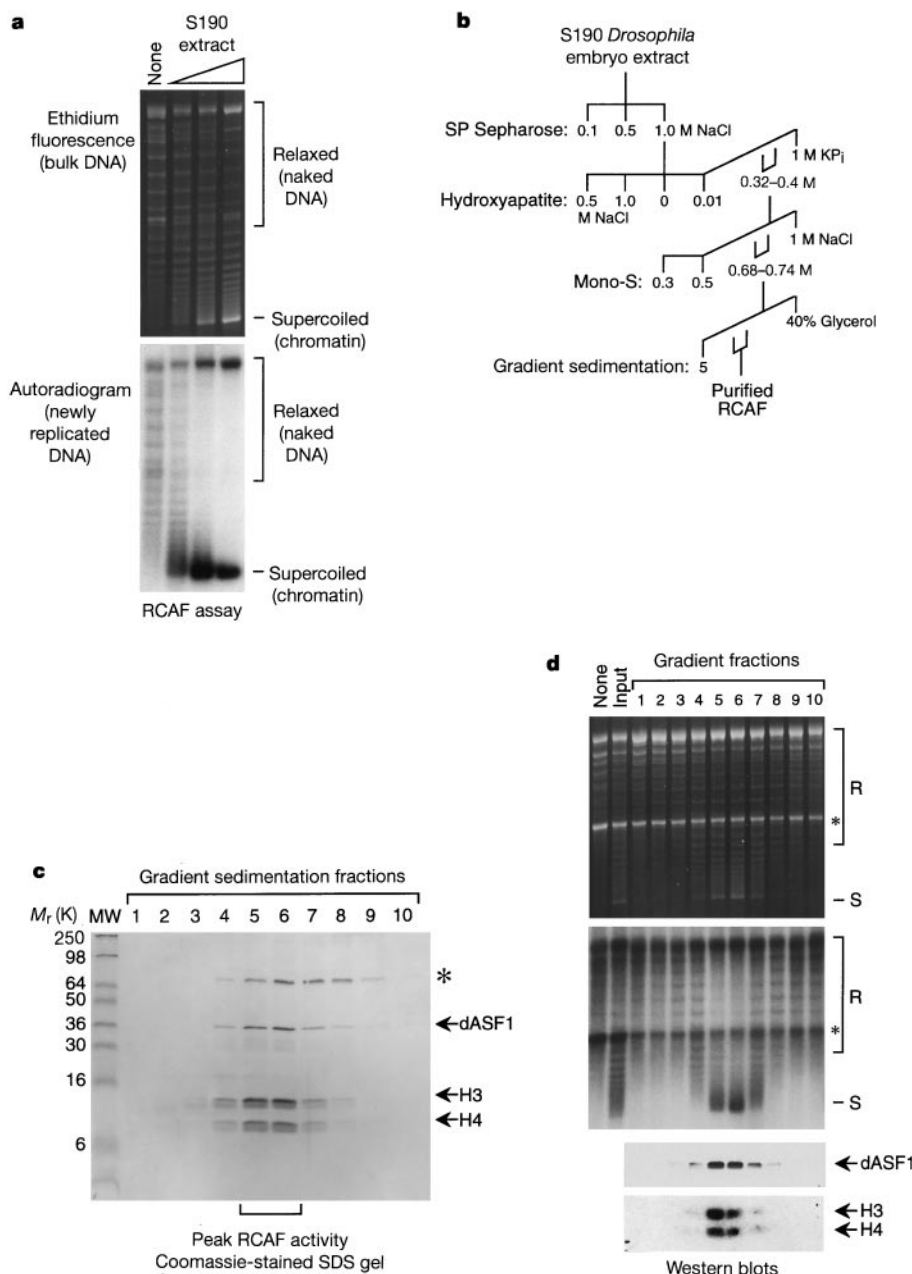


Figure 1 Purification of RCAF activity. **a**, Addition of the *Drosophila* S190 extract restores replication-coupled chromatin assembly, as determined by the RCAF assay. The asterisks denote a linear DNA species generated by a nuclease activity in the replication extract. **b**, Scheme for the purification of RCAF. **c**, RCAF is a complex of three polypeptides. Gradient fractions from the sedimentation analysis were subjected to SDS-PAGE (15% gel) and staining with Coomassie blue. The asterisk denotes a polypeptide whose elution is not correlated with RCAF activity and is assumed to be a contaminating species.

d, RCAF assay of the sedimentation analysis fractions shown in **c**. The same fractions were subjected to western blot analysis with an affinity-purified polyclonal antiserum raised against full-length recombinant dASF1, polyclonal antiserum recognizing *Drosophila* H3 (Mark Levenstein, University of California, San Diego) and polyclonal antiserum recognizing acetylated H4 (C. David Allis, University of Virginia). RCAF sedimented with an apparent relative molecular mass of 66K, as determined by comparison with size standards analysed in parallel (data not shown).

sequences; GenBank accession numbers AI073760 and AA769906). ASF1 exhibits no apparent similarity to other proteins involved in chromatin assembly.

Quantitative western blot analyses of dASF1 throughout *Drosophila* development revealed that it is highly abundant in embryos (at ~200,000 molecules per cell) and less abundant in larvae, pupae and adults (data not shown). For comparison, the p55 subunit of dCAF-1, which is also a component of chromatin remodelling, histone deacetylase and histone acetylase complexes, is present at ~220,000 molecules per cell¹⁰. The presence of the highest levels of dASF1 in embryos is correlated with the large amounts of DNA replication and chromatin assembly that occur during early embryogenesis.

Amino-terminal amino acid sequencing of the smaller RCAF subunits revealed that they are *Drosophila* histones H3 and H4. Furthermore, this microsequencing analysis indicated that the Lys 5 and Lys 12 of RCAF H4 were acetylated, whereas Lys 8 was not detectably acetylated. Similarly, Lys 14 of RCAF H3 was fully acetylated, whereas Lys 4 and Lys 9 were not detectably acetylated. Significantly, the specific patterns of lysine acetylation present in the H3 and H4 components of RCAF (AcK14 of H3, and AcK5 and AcK12 of H4) are identical to those of newly synthesized *Drosophila* histones H3 and H4¹⁶. Newly synthesized histones H3 and H4 are modified by acetylation at these specific lysine residues before their incorporation into chromatin and then deacetylated shortly after their deposition on to DNA.

The function of ASF1 as a chaperone for newly synthesized histones during S phase is consistent with the expression of yeast ASF1 in late G₁ immediately before histone synthesis. Furthermore, the ASF1 gene is coordinately regulated with genes encoding DNA replication and repair proteins, owing to the presence of *cis*-acting promoter elements, termed *Mlu*I cell cycle boxes in the promoter of ASF1 and more than 20 *S. cerevisiae* genes involved in DNA

replication and DNA repair^{6,15,17}.

Further verification that dASF1, H3 and H4 are in the RCAF complex was provided by the strong correlation between the migration of the RCAF activity and the dASF1, H3 and H4 polypeptides, as determined by western blotting, through the last three steps for the purification procedure (Fig. 1d, and data not shown). Taken together, these results confirm that RCAF consists of a complex of dASF1 and specifically acetylated histones H3 and H4.

To address the involvement of dASF1 in RCAF-mediated chromatin assembly, we tested various sources of purified histones for their ability to mediate chromatin assembly in the absence of dASF1. Hyperacetylated core histones, bulk core histones or recombinant H3–H4 failed to exhibit any detectable RCAF activity, even when the histones were included in much greater quantities than those present in RCAF (Fig. 2a). These results indicate that dASF1 is essential for the chromatin assembly activity of the RCAF complex.

Because RCAF and CAF-1 both act as chaperones for histones H3–H4¹⁸, we investigated whether RCAF functions in an additive or a synergistic manner with CAF-1. We therefore performed RCAF assays with purified RCAF alone, CAF-1 alone, or both RCAF and CAF-1 together (Fig. 2b). These experiments revealed that RCAF together with CAF-1 assembled newly replicated DNA into chromatin with a greater efficiency than that achieved with either four times the same amount of RCAF alone or twice the same amount of CAF-1 alone, in a T-antigen-dependent manner (Fig. 2b, and data not shown). These results indicate that RCAF and CAF-1 can function synergistically in the assembly of chromatin on to newly synthesized DNA.

As a complement to the biochemical studies, we investigated the function of RCAF and CAF-1 *in vivo* in *S. cerevisiae*. For RCAF we analysed a disruption of the ASF1 gene⁶, whereas for CAF-1 we

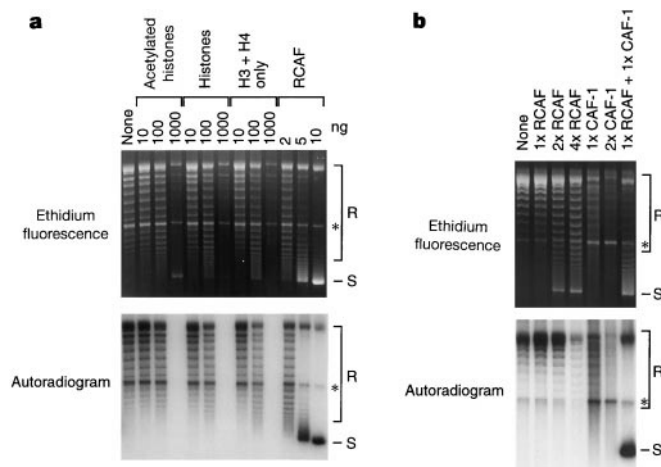


Figure 2 RCAF cannot be replaced by histones and acts synergistically with CAF-1 in the assembly of chromatin on to newly replicated DNA. **a**, Histones cannot substitute for RCAF activity. Hyperacetylated HeLa core histones (Acetylated histones), HeLa core histones (Histones), recombinant *Drosophila* H3–H4 tetramers or purified RCAF were assayed for their relative abilities to mediate chromatin assembly on to newly replicated DNA (the amount of H3–H4 tetramers (ng) in each reaction is indicated). dCAF-1 was included in all reactions. **b**, Synergism between RCAF and CAF-1 in chromatin assembly. The standard quantity of RCAF or CAF-1 that mediates chromatin assembly is denoted by 1x. 2x and 4x indicate two and four times the standard quantities of factors. Supercoiling of bulk DNA was seen with larger quantities of RCAF in the absence of exogenous CAF-1 or in the absence of DNA replication (data not shown). These findings indicate that CAF-1 might render the newly replicated DNA a more favourable substrate for chromatin assembly by RCAF. The asterisks denote a linear DNA species generated by a nuclease activity in the replication extract.

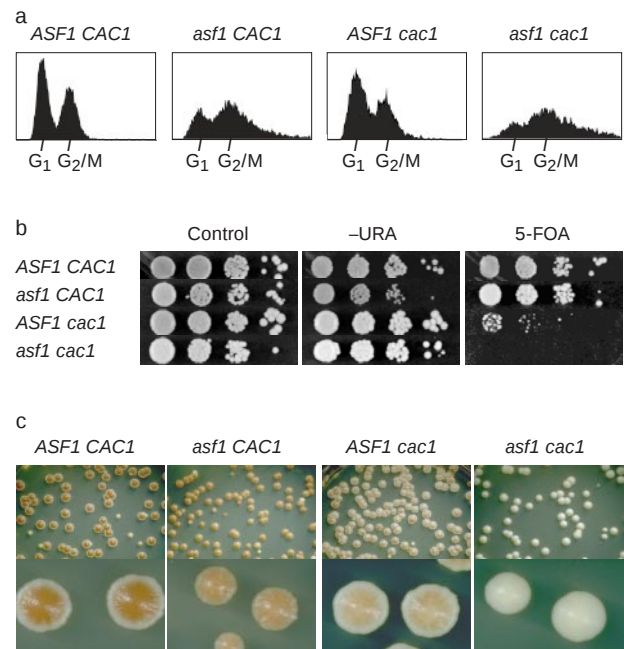


Figure 3 Defects in growth and transcriptional silencing after the loss of both RCAF and CAF-1 activity. **a**, Accumulation of *asf1* and *asf1 cac1* mutant yeast strains in G₂/M. The DNA content of asynchronous cultures of strains ROY 1172 (*ASF1 CAC1*), ROY 1170 (*asf1 CAC1*), ROY 1171 (*ASF1 cac1*) and ROY 1169 (*asf1 cac1*) were determined by flow cytometric analysis. **b**, Telomeric silencing of *URA3*. Strain ROY 1172 (*ASF1 CAC1*), ROY 1169 (*asf1 cac1*), ROY 1170 (*asf1 CAC1*) and ROY 1171 (*ASF1 cac1*) were serially diluted on control plates, plates lacking uracil (–URA) and plates containing 5-FOA. **c**, *HMR* silencing of *ADE2*. Two different magnifications of the colonies for strains ROY 1172 (*ASF1 CAC1*), ROY 1178 (*asf1 CAC1*), ROY 1179 (*ASF1 cac1*) and ROY 1177 (*asf1 cac1*) are shown.

analysed a disruption of *CAC1*, which encodes the largest subunit of CAF-1 (disruption of *CAC1* results in the same phenotypes as disruption of all three genes encoding CAF-1 subunits)¹⁹.

We first examined the effects of disruptions in *ASF1* and *CAC1* on growth. It had been seen previously that *asf1* mutant yeast strains grow poorly^{6,14}, whereas *cac1* mutants are not defective for growth¹⁹. The estimated doubling times for cultures of the *ASF1 CAC1* control, the *asf1* mutant, the *cac1* mutant and the *asf1 cac1* mutant were 2.3, 3.1, 1.8 and 3.7 h respectively, indicating that the *asf1 cac1* double mutant grows even more slowly than the *asf1* mutant.

To investigate which phases of the cell cycle were lengthened by disruption of *ASF1*, we performed a flow cytometric analysis (Fig. 3a). We have found that asynchronous cultures of the *asf1* mutant yeast accumulated with a DNA content that corresponded to the G₂/M phases of the cell cycle. In addition the *asf1* mutant cells seemed to remain in S phase longer than the *ASF1 CAC1* control strain (Fig. 3a). The cells of the *asf1 cac1* mutant culture were also markedly skewed towards high DNA content (Fig. 3a). The broadened flow cytometric profiles for the *asf1* and *asf1 cac1* mutants were

correlated with the wide range of cell sizes of these strains, with ~10% of the population having twice the normal cell diameter (data not shown). Also, the *asf1* mutant exhibited a temperature-sensitive phenotype that was enhanced in the *asf1 cac1* mutant (data not shown). These results indicate that *ASF1* is essential for normal progression through the cell cycle. Moreover, in the absence of RCAF function, the loss of CAF-1 leads to extra growth defects. These findings are consistent with a partly redundant function of RCAF and CAF-1 in chromatin assembly, which is essential for cell growth²⁰. But the *asf1 cac1* double mutant was still viable, and hence one or more other histone chaperones with related functions might exist.

Next, to test for the integrity of repressive chromatin structures, we examined the effects of disruptions in *ASF1* and *CAC1* on transcriptional silencing. To examine telomeric silencing, we used a yeast strain containing a *URA3* gene inserted next to a telomere. In this strain, derepression of *URA3* leads to cell death on media containing 5-fluoro-orotic acid (5-FOA)²¹. With this assay, *cac1* mutants displayed partial derepression of *URA3* silencing^{19,22} (Fig. 3b). We observed no significant derepression of telomeric

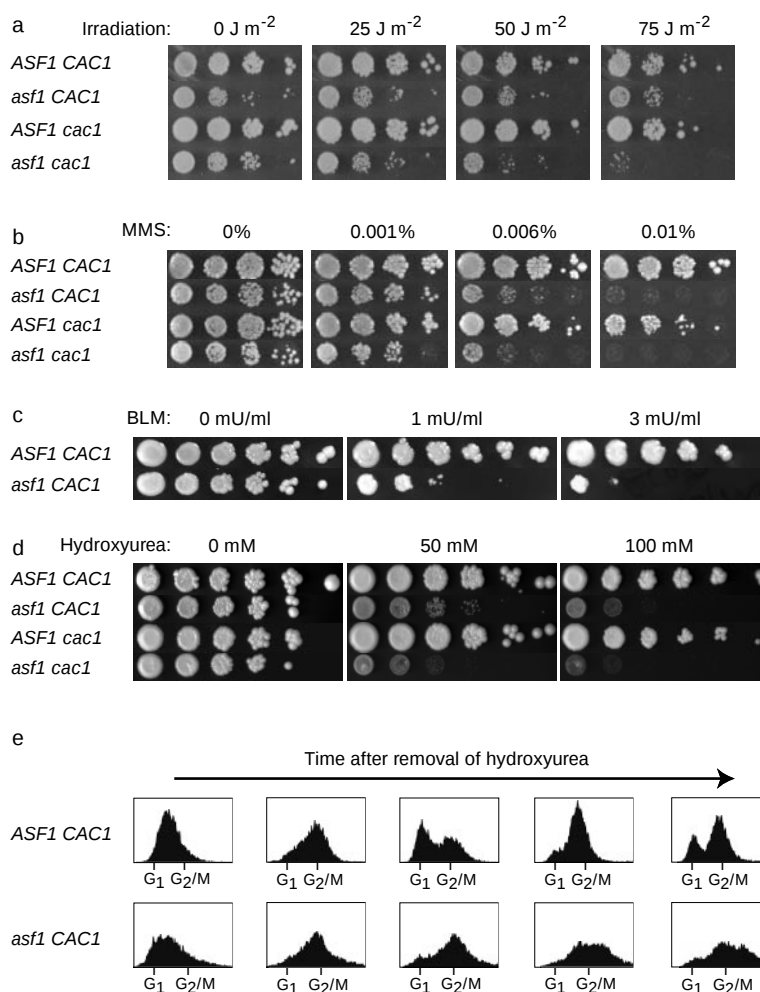


Figure 4 Effects of disruptions to *ASF1* and *CAC1* on sensitivity to mutagens. **a**, *asf1 cac1* double mutants are more sensitive to ultraviolet irradiation than either single mutant. Strains ROY 1172 (*ASF1 CAC1*), ROY 1170 (*asf1 CAC1*), ROY 1171 (*ASF1 cac1*) and ROY 1169 (*asf1 cac1*) were exposed to the indicated doses of ultraviolet radiation. **b**, *asf1* mutants, but not *cac1* mutants, exhibit sensitivity to MMS. Serial dilutions of strains JRY 2334 (*ASF1 CAC1*), RS 1453 (*asf1 CAC1*), ROY 1181 (*ASF1 cac1*) and ROY 1180 (*asf1 cac1*) were spotted on to YPD plates containing the indicated concentrations of MMS. **c**, The *asf1* mutant exhibits sensitivity to BLM. Strains ROY 1172 (*ASF1 CAC1*), ROY 1170 (*asf1 CAC1*), ROY 1171 (*ASF1 cac1*) and ROY 1169 (*asf1 cac1*) were spotted on to YPD

plates containing the indicated concentrations of BLM. **d**, *asf1* mutants, but not *cac1* mutants, exhibit sensitivity to hydroxyurea. Strains ROY 1172 (*ASF1 CAC1*), ROY 1170 (*asf1 CAC1*), ROY 1171 (*ASF1 cac1*) and ROY 1169 (*asf1 cac1*) were spotted on to YPD plates containing the indicated concentrations of hydroxyurea. **e**, *asf1* mutant yeast arrest in G₂/M after release from hydroxyurea-induced arrest. Aliquots of cultures of strains ROY 1172 (*ASF1 CAC1*), ROY 1170 (*asf1 CAC1*), ROY 1171 (*ASF1 cac1*) and ROY 1169 (*asf1 cac1*) were taken at various times after removal of hydroxyurea (200 mM) for flow cytometric analysis.

silencing of *URA3* in the *asf1* mutant strain when compared with the *ASF1 CAC1* control strain. Notably, *URA3* was completely derepressed in the *asf1 cac1* double mutant, which indicates that the *asf1 cac1* strain had greater defects in the assembly of silenced chromatin than either of the single-mutant strains.

We also tested whether the *asf1 cac1* mutant was defective for silencing at the *HMR* and *HML* mating-type loci. We used the *HMR::ADE2* reporter system to measure the expression of an *ADE2* gene inserted at the *HMR* mating-type locus. In this assay, silencing of *ADE2* yields red colonies, whereas derepression of *ADE2* yields white colonies. The *cac1* mutant yielded pink colonies, indicating slight derepression of *ADE2* (Fig. 3c), whereas the *asf1* mutant had no significant defect in the silencing of *ADE2* at *HMR*. But the *asf1 cac1* double mutant displayed complete derepression of *ADE2*, as shown by the white colonies (Fig. 3c). Comparable results were obtained for the silencing of a *TRP1* gene inserted at the *HML* locus (data not shown). In addition, the *asf1 cac1* double mutant failed to mate in a mating assay (data not shown). Mutation of *CAC1* alone therefore has a more pronounced effect on telomeric and mating-type silencing than mutation of *ASF1* alone. But the disruption of both *CAC1* and *ASF1* results in a more drastic defect in the integrity of silenced chromatin. These results suggest that RCAF and CAF-1 each participate in the assembly of silenced chromatin.

Next, we investigated whether RCAF is involved in chromatin assembly during DNA repair. We examined the effects of disruptions in *ASF1* and *CAC1* on sensitivity to ultraviolet and found that the *asf1 cac1* double mutant was significantly more sensitive to ultraviolet irradiation than either single-mutant strain alone (Fig. 4a). These results suggest that RCAF and CAF-1 are each involved in the assembly of chromatin in the course of repair of ultraviolet-irradiated DNA.

We also investigated whether CAF1 and RCAF function in other DNA repair pathways. It had been shown previously that *asf1* mutants are sensitive to the DNA damage-inducing mutagen methyl methane sulphonate (MMS)⁶, which creates adducts and apurinic sites that become single- and double-strand DNA breaks²³. The *asf1* mutant, but not the *cac1* mutant, exhibited sensitivity to MMS (Fig. 4b). Moreover, the *asf1 cac1* double mutant displayed a similar degree of sensitivity to MMS to that of the *asf1* single mutant (Fig. 4b). These results suggest that RCAF, but not CAF-1, is involved in the repair of MMS-induced DNA damage. Similarly, the *asf1* mutant was sensitive to bleomycin (BLM), a radiomimetic drug that creates double-strand DNA breaks by a concerted free-radical attack on sugar moieties²⁴ (Fig. 4c).

To investigate where MMS and BLM block the cell cycle in the *asf1* mutant, we performed flow cytometric analysis of the drug-treated yeast cultures. We found that treatment with MMS or BLM arrested the *asf1* mutant and the *ASF1 CAC1* control strains in the G_2/M phases of the cell cycle. But the *asf1* mutant failed to proceed through the cell cycle after removal of the drug, instead remaining in G_2/M (data not shown). These results suggest that RCAF is involved in the repair of double-strand DNA breaks, presumably in G_2 , through the assembly of newly repaired DNA into chromatin. Importantly, these findings provide evidence that chromatin assembly *de novo* is integral to double-strand DNA break repair.

To investigate the relationship between DNA synthesis and RCAF further, we used hydroxyurea, which inhibits DNA synthesis by depleting the nucleotide pool²⁵. Treatment with hydroxyurea causes incomplete duplication of the chromosomes, which is then completed in G_2 by double-strand DNA repair mechanisms. We found that the *asf1* mutant, but not the *cac1* mutant, was sensitive to hydroxyurea (Fig. 4d). Treatment with hydroxyurea arrested *asf1* and the *ASF1 CAC1* control strains in S phase (Fig. 4e). But after removal of hydroxyurea, the *ASF1 CAC1* control strain was able to proceed through the cell cycle, whereas the majority of the *asf1*

mutant cells progressed into G_2/M and then did not proceed further. These findings suggest that in the absence of *ASF1/RCAF* there is insufficient chromatin assembly activity after hydroxyurea-mediated arrest and subsequent release to generate chromosomes that are competent to proceed through G_2/M .

Significantly, yeast cells that are depleted of histone H4 or histone H2B arrest in G_2/M and are unable to complete chromosomal segregation^{26,27}. Therefore the accumulation of *asf1* mutant yeast in G_2/M after treatments that inhibit DNA replication or damage DNA (Fig. 4) are consistent with a role for RCAF during the assembly of chromatin on to newly synthesized DNA *in vivo*.

We have identified and characterized a novel factor, RCAF, that mediates the assembly of newly synthesized DNA into chromatin. The results suggest that RCAF and CAF-1 can function coordinately in replication-coupled chromatin assembly, but also that each factor is able to effect some chromatin assembly in the absence of the other. Furthermore, it seems that RCAF mediates chromatin assembly after the repair of double-strand DNA breaks. These findings should provide a better understanding of the chromatin assembly process and lead to new insight into the relation between DNA replication and gene expression. In addition, further analysis of RCAF should be relevant to the understanding of conditions such as the predisposition to cancer and neurological abnormalities that arise from defective DNA repair. □

Methods

RCAF assay

SV40 DNA replication reactions were performed and analysed as described previously¹⁰, with the following modifications. The SV40-origin-containing plasmid pSVL- Ψ CRK was replicated by the addition of SV40 T-antigen and a replication extract from human 293 cells. The amount of template DNA included was determined empirically for each extract as the amount of DNA that inhibited chromatin assembly. The inclusion of [α -³²P]dATP results in labelling of the newly replicated DNA. All RCAF assay reactions included *Drosophila* CAF-1 that had been purified through the hydroxyapatite column¹⁰. The reactions were deproteinized and the DNA was resolved by 1% agarose-gel electrophoresis. The gels were stained with ethidium bromide to reveal the bulk plasmid DNA and subjected to autoradiography to reveal the newly replicated DNA.

Purification of RCAF

RCAF was purified from supernatants from 190,000g centrifugations (S190) that were prepared from *Drosophila* embryos collected from 0 to 6 h after egg deposition²⁸. The RCAF activity was purified through four successive purification steps (outlined in Fig. 1b).

Protein sequencing of RCAF

The following peptide information was obtained for the dASF1 subunit of RCAF by using protein microsequencing techniques described previously¹⁰; MIYVGSASEEHDQVLDTI, YVGPVPEGRH, PLFEK, PRVTRFK, INWDYGHINGNGVENGHQD. N-terminal sequences for H3 (ARTKQTARKSTGGKAPR) and for H4 (GRGKGKGLGKGG) were obtained by automated Edman degradation. The N-terminal acetylation of RCAF H4 was deblocked as described¹⁶.

Cloning of dASF1

Degenerate primers corresponding to the expected coding sequence were used to generate a partial cDNA fragment by PCR for the dASF1 component of RCAF. Ten independent cDNAs were isolated by screening a *Drosophila* embryo cDNA library (0–4 h embryos) in λ ZAPII (Stratagene) with the radiolabelled cDNA fragment. By hybridization *in situ* to *Drosophila* polytene chromosomes, the dASF1 locus was mapped to region 76C1–3.

Yeast analyses

For all yeast analyses, two independent sets of single yeast colonies of the appropriate strains were grown in YPD medium overnight at 30 °C, diluted to an initial absorbance of 0.05–0.1 at 600 nm before further analysis. For the analysis of telomeric silencing, tenfold serial dilutions of the appropriate strains were spotted on to YMD–TRP, YMD–TRP–URA and 1 mg ml^{−1} 5-FOA–TRP plates. For the analysis of sensitivity to ultraviolet, serial dilutions were spotted on to YPD plates and irradiated by using a Stratlinker. All plates were grown at 30 °C for 2 to 3 d before photography.

Cells for fluorescence-activated cell sorting (FACS) were stained with propidium iodide, sonicated and analysed on a FACScan flow cytometer (Becton-Dickinson) with LysisII Software. Recovery from a 3 h treatment with hydroxyurea was determined by washing the cells with YPD medium and growing at 30 °C in YPD. Samples were taken for FACS analysis at 0 min, 45 min, 1 h 30 min, 2 h 15 min and 3 h after the removal of hydroxyurea. All experiments presented were performed at least twice independently (up to four times in some cases) with reproducible results.

Yeast strains used

Strains and genotypes were as follows: ROY 1169, *MAT α ade2-1 LYS2 leu2-3,112 his3-11 trp1-1 ura3-1 cac1::LEU2 asf1::his5⁺ HMR::ADE2 TELVIL::URA3*; ROY 1170, *MAT α ade2-1 LYS2 leu2-3,112 his3-11 trp1-1 ura3-1 asf1::his5⁺ HMR::ADE2 TELVIII::URA3*; ROY 1171, *MAT α ade2-1 LYS2 leu2-3,112 his3-11 trp1-1 ura3-1 cac1::LEU2 HMR::ADE2 TELVIII::URA3*; ROY 1172, *MAT α ade2-1 LYS2 leu2-3,112 his3-11 trp1-1 ura3-1 HMR::ADE2 TELVIII::URA3*; ROY 1177, *MAT α ade2-1 LYS2 leu2-3,112 his3-11 trp1-1 ura3-1 cac1::LEU2 asf1::his5⁺ HMR::ADE2 TELVIII::URA3*; ROY 1178, *MAT α ade2-1 LYS2 leu2-3,112 his3-11 trp1-1 ura3-1 asf1::his5⁺ HMR::ADE2 TELVIII::URA3*; ROY 1179, *MAT α ade2-1 LYS2 leu2-3,112 his3-11 trp1-1 ura3-1 cac1::LEU2 HMR::ADE2 TELVIII::URA3*; ROY 1180, *MAT α ade2-1 LYS2 leu2-3,112 his3-11 trp1-1 ura3-1 asf1::his5⁺ cac1::LEU2*; ROY 1181, *MAT α ade2-1 LYS2 leu2-3,112 his3-11 trp1-1 ura3-1 cac1::LEU2*; RS 1453, *MAT α ade2-1 LYS2 leu2-3,112 his3-11 trp1-1 ura3-1 asf1::his5⁺*; JRY 2334, *MAT α ade2-1 LYS2 leu2-3,112 his3-11 trp1-1 ura3-1*. All the strains are isogenic with W-303.

Received 9 August; accepted 29 October 1999.

- Grunstein, M. Histone acetylation in chromatin structure and transcription. *Nature* **389**, 349–352 (1997).
- Kornberg, R. D. & Lorch, Y. Chromatin-modifying and -remodelling complexes. *Curr. Opin. Genet. Dev.* **9**, 148–151 (1999).
- Ito, T., Tyler, J. K. & Kadonaga, J. T. Chromatin assembly factors: a dual function in nucleosome formation and mobilization? *Genes Cells* **2**, 593–600 (1997).
- Adams, C. R. & Kamakaka, R. T. Chromatin assembly: biochemical identities and genetic redundancy. *Curr. Opin. Genet. Dev.* **9**, 185–190 (1999).
- Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. R. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**, 251–260 (1997).
- Le, S., Davis, C., Konopka, J. B. & Sternglanz, R. Two new S-phase specific genes from *Saccharomyces cerevisiae*. *Yeast* **13**, 1029–1042 (1997).
- Stillman, B. W. & Gluzman, Y. Replication and supercoiling of simian virus 40 DNA in cell extracts from human cells. *Mol. Cell. Biol.* **5**, 2051–2060 (1985).
- Stillman, B. Chromatin assembly during SV40 DNA replication *in vitro*. *Cell* **45**, 555–565 (1986).
- Smith, S. & Stillman, B. Purification and characterization of CAF-I, a human cell factor required for chromatin assembly during DNA replication *in vitro*. *Cell* **58**, 15–25 (1989).
- Tyler, J. K., Bulger, M., Kamakaka, R. T., Kobayashi, R. & Kadonaga, J. T. The p55 subunit of *Drosophila* chromatin assembly factor 1 is homologous to a histone deacetylase-associated protein. *Mol. Cell. Biol.* **16**, 6149–6159 (1996).
- Ito, T., Bulger, M., Pazin, M. J., Kobayashi, R. & Kadonaga, J. T. ACF, an ISWI-containing and ATP-utilizing chromatin assembly and remodeling factor. *Cell* **90**, 145–155 (1997).
- Ito, T., Bulger, M., Kobayashi, R. & Kadonaga, J. T. *Drosophila* NAP-1 is a core histone chaperone that functions in ATP-facilitated assembly of regularly spaced nucleosomal arrays. *Mol. Cell. Biol.* **16**, 3112–3124 (1996).
- Ito, T., Tyler, J. K., Bulger, M., Kobayashi, R. & Kadonaga, J. T. ATP-facilitated chromatin assembly with a nucleoplasmid-like protein from *Drosophila melanogaster*. *J. Biol. Chem.* **271**, 25041–25048 (1996).
- Singer, M. S. *et al.* Identification of high-copy disruptors of telomeric silencing in *Saccharomyces cerevisiae*. *Genetics* **150**, 613–632 (1998).
- Spellman, P. K. *et al.* Comprehensive identification of cell cycle-regulated genes of the yeast

- Saccharomyces cerevisiae* by microarray hybridization. *Mol. Biol. Cell* **9**, 3273–3287 (1998).
- Sobel, R. E., Cook, R. G., Perry, C. A., Annunziato, A. T. & Allis, C. D. Conservation of deposition-related acetylation sites in newly synthesized histones H3 and H4. *Proc. Natl Acad. Sci. USA* **92**, 1237–1241 (1995).
- McIntosh, E. M. MCB elements and the regulation of DNA replication genes in yeast. *Curr. Genet.* **24**, 185–192 (1993).
- Verreault, A., Kaufman, P. D., Kobayashi, R. & Stillman, B. Nucleosome assembly by a complex of CAF-I and acetylated histones H3/H4. *Cell* **87**, 95–104 (1996).
- Kaufman, P. D., Kobayashi, R. & Stillman, B. Ultraviolet radiation sensitivity and reduction of telomeric silencing in *Saccharomyces cerevisiae* cells lacking chromatin assembly factor-I. *Genes Dev.* **11**, 345–357 (1997).
- Ling, X., Harkness, T. A. A., Schultz, M. C., Fisher-Adams, G. & Grunstein, M. Yeast histone H3 and H4 amino termini are important for nucleosome assembly *in vivo* and *in vitro*: redundant and position-independent functions in assembly but not in gene regulation. *Genes Dev.* **10**, 686–699 (1996).
- Gottschling, D. E., Aparicio, O. M., Billington, B. L. & Zakian, V. A. Position effect at *S. cerevisiae* telomeres: reversible repression of Pol II transcription. *Cell* **63**, 751–762 (1990).
- Monson, E. K., Bruin, D. de & Zakian, V. A. The yeast *cac 1* protein is required for the stable inheritance of transcriptionally repressed chromatin at telomeres. *Proc. Natl Acad. Sci. USA* **94**, 13081–13086 (1997).
- Schwartz, J. L. Monofunctional alkylating agent-induced S-phase-dependent DNA damage. *Mutat. Res.* **216**, 111–118 (1989).
- Povirk, L. F. DNA damage and mutagenesis by radiomimetic DNA-cleaving agents: bleomycin, neocarzinostatin and other enediynes. *Mutat. Res.* **355**, 71–89 (1996).
- Elledge, S. J. Cell cycle checkpoints: preventing an identity crisis. *Science* **274**, 1664–1672 (1996).
- Han, M., Chang, M., Kim, U.-J. & Grunstein, M. Histone H2B repression causes cell-cycle-specific arrest in yeast: effects on chromosomal segregation, replication, and transcription. *Cell* **48**, 589–597 (1987).
- Kim, U.-J., Han, M., Kayne, P. & Grunstein, M. Effects of histone H4 deletion on the cell cycle and transcription of *Saccharomyces cerevisiae*. *EMBO J.* **7**, 2211–2219 (1988).
- Kamakaka, R. T., Bulger, M. & Kadonaga, J. T. Potentiation of RNA polymerase II transcription by Gal4-VP16 during but not after DNA replication and chromatin assembly. *Genes Dev.* **7**, 1779–1795 (1993).

Acknowledgements

We thank P. Willy, D. Fyodorov, X. Huang, J. Butler, M. Levenstein, N. Dhillon, D. Donze, S. Ghidelli, L. Krushel and E. Stone for critical reading of the manuscript; L. Pillus for helpful suggestions and discussions; T. Lavery and G. Rubin for determination of the cytological location of *dASF1*; K. Collins for generating recombinant *dASF1* protein; M. Levenstein for H3 antibodies and purified histones H3–H4; D. Allis for H4 antibodies; and R. Sternglanz, P. Kaufman, D. Shore, D. Gottschling and J. Rine for providing various strains and plasmids used in this study. This work was supported by grants from the NIH to J.T.K., R.K. and R.T.K.; J.K.T. is a Leukemia Society of America Special Fellow.

Correspondence and requests for materials should be addressed to J.T.K. (e-mail: jkadonaga@ucsd.edu). The GenBank accession number for the *dASF1* cDNA is AF189278.

A trick of the light

Lasers for tissue processing, imaging, measurement, ablation and so on.

DUO dye lasers

Laser Science, Inc. www.laserscience.com

Simultaneous UV and visible laser light

DUO-220 and DUO-221 are grating-based dye lasers tunable with a micrometer from 360 to 700 nm and 600 to 950 nm respectively. Third in the series, the DUO-210 is of a simpler design, for higher energy output at a fixed wavelength. All DUO lasers have dual-beam exit ports for dye and nitrogen laser output. A six-position selector wheel allows users to select dye laser output, reflect out an unaltered nitrogen beam at 337 nm or add a beamsplitter for simultaneous visible and UV output.

Reader Service No. 100

FLA-3000G

Fuji Photo Film Co. www.fujifilm.com

Three lasers combine with a large scanner for fluorescent and radioisotope imaging

Equipped with two Second Harmonic Generation lasers (blue/473 nm, green/532 nm) and a He-Ne laser (red/633 nm), the FLA 3000G can process samples labelled with either radioisotopes or fluorescent dyes. The addition of the SHG green/532 nm laser makes it possible to discriminate between Cy2 and Cy5, as well as Cy3 and Cy5 labelled samples. Used with a wide-format imaging plate, the FLA-3000G is suitable for gene expression profiling, using commercially available cDNA membranes.

Reader Service No. 101

S2000-1550 Spectrometer

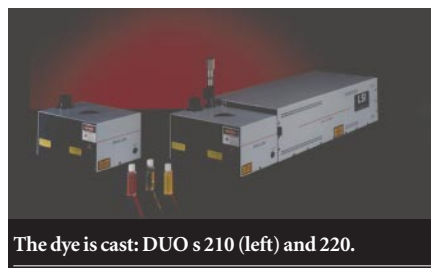
Ocean Optics Inc. www.oceanoptics.com

NIR spectrometry for laser characterization

This PC-based UV/VIS miniature fibre optic spectrometer has been reengineered as an NIR-sensitive, high-resolution system for laser and laser diode wavelength characterization. The S2000-1550 has a silicon linear CCD array with a special coating that extends the silicon array's responsivity into the NIR, from 1,450 to 1,650 nm. This makes the S2000-



Character role: the S2000-1550 spectrometer.



The dye is cast: DUO s 210 (left) and 220.

1550 especially useful for analysing the real-time spectral output of various lasers and laser diodes, including both pulsed and continuous-wave lasers.

Reader Service No. 102

LabRam Infinity dual laser

Instruments SA (UK) www.isa-gs.co.uk

Flexible dual laser for specific analysis

Following the introduction of the LabRam Infinity analytical Raman microscope, Instruments SA has launched a dual-laser model for analytical and QC/QA Raman analysis. Developed as a low-cost system for the highly specific analysis of solid, liquid and solution samples, it offers fast routine analysis and allows easy choice of the laser source most appropriate to reduce problems such as sample heating or fluorescence.

Reader Service No. 103

GeoLas

MicroLas www.lambdaphysik.com

Short wavelength to reduce sample-heating

GeoLas is an excimer laser ablation system for Inductively Coupled Plasma Mass Spectrometry. Introduced by MicroLas in affiliation with Lambda Physik and JENOPTIK, it is optimized for analysis of geological samples and novel materials, such as superconducting ceramics and glasses. GeoLas is built around a Lambda Physik COMPex 100 excimer laser operating with argon fluoride. The 193 nm output can ablate transparent materials difficult to process at longer wavelengths. The 193 nm photons are strongly absorbed by quartz, fluorites, apatites, calcite, garnets and refractory igneous materials such as olivine.

Reader Service No. 104

VG Microprobe II

VG Elemental www.vgelemental.com

Advanced ultraviolet laser ablation

This is an advanced version of the VG Microprobe, the first commercial UV laser ablation system to be fully integrated into ICP-MS hardware and software. The new system uses a

frequency quadrupled Nd:YAG laser, and has a spectrally pure 266 nm beam with a PC-controlled variable energy of up to 4 mJ at the sample surface. Varying the ablation frequency with Q-switch gating allows the laser always to fire at its maximum flash lamp voltage to optimize stability.

Reader Service No. 105

LSM 5 PASCAL

Carl Zeiss www.zeiss.de

A budget-priced laser scanning microscope

The LSM 5 PASCAL system, aimed at the single user and small groups, can be configured for both biomedical, and materials research and testing uses. Many of the features of the high-end LSM 510 are retained in the LSM 5 PASCAL: scanning fields of unparalleled size, single images with more than 4 million pixels in up to 4,096 grey levels, and high flexibility in selecting the scanning modes and powerful, user-friendly software.

Reader Service No. 106

Visualize

Applied Scintillation Technology

www.appscintech.com

The spectrum of laser light made visible

This range — VisualizeIR, VISualize and UVisualize — allows visualization of lasers emitting light across the spectrum. It includes optical bench- and rail-mounted devices as well as wand and 'credit card' formats. They don't need to be pre-charged, offer non-fade IR detection, non-reflective encapsulation and extended afterglow for viewing pulsed UV lasers. Optical bench- and rail-mounted versions are available in standard 40 mm size, and the wand has a demountable 25 mm disc.

Reader Service No. 107



Good looking: visualize laser light with Visualize.

OPTex Excimer Laser

Lambda Physik www.lambdaphysik.com

Portable excimer laser extends its range

The compact OPTex excimer laser is now available in 157 nm specification, in addition to the existing range of 193, 248, 308 and 351 nm. The OPTex Fluorine laser tested pulse energies of 1 mJ at 5 Hz, repetition rates up to 200 Hz and average power at 180 mW. The NovaTube design is claimed to eliminate gas corrosion and contamination, while increasing the tube's lifetime and reducing operating costs. The OPTex comes equipped for operating in the UV range, and includes Windows-ware. Applications include laser microsurgery, spectroscopy, bioanalysis, photochemistry, ablating materials and ionization.

Reader Service No. 108



On the move: excimer portability from OPTex.

200-kW peak power and pulse energies of >55µJ. UDL dye lasers, designed to attach to the GNL series, are available in three versions depending on linewidth requirements.

Reader Service No. 110

Lumesphere Plus

Optronic Laboratories www.isa-gs.co.uk

Precise measurement with a total luminous flux system

The Lumesphere Plus is an integrated turnkey system for automated total luminous flux measurements. It has been designed to measure the power of various lamps and light sources precisely, in accordance with CIE and National Institute of Standards and Technology recommendations. With dedicated software, the system is precise enough for researchers, yet simple enough to be operated by inexperienced personnel.

Reader Service No. 111

PC Laser Beam Analyser

Spiricon www.spiricon.com

High resolution of a laser beam profile

The LBA-500PC version 2.60 laser beam analyser incorporates a number of features for enhanced performance: 10-bit and 12-bit intensity resolution of the profile of a laser beam is now possible, with spatial resolutions of up to 1,024 × 1,024 pixels. Also included in this release is the ability to export files compatible for importation into Frequency

Resolved Optical Gating (FROG) software. Other improvements include a 2 point far field divergence method, a column and row summation export file, and ability to copy images and results to the Windows clipboard.

Reader Service No. 112

PALM Robot-MicroBeam

PALM Mikrolaser Technologie

www.palm-mikrolaser.com

Tissue processing by laser

The Robot-MicroBeam combines a UV laser microbeam with a motorized, computer-controlled microscope stage and a micro-manipulator (Robot-Stage and Robot-Manipulator). The makers promise that cutting and catapulting is as easy as a computer game and takes just minutes to learn. Using the combined laser microbeam dissection and laser pressure catapulting techniques, single cells can be processed in a few seconds and tissue areas containing thousands of cells take less than 10 minutes. DNA, RNA and protein recovery are entirely preserved.

Reader Service No. 113

Vitrea

Vital Images www.vitalimages.com

Laser visualization for medical diagnosis

The Vitrea medical visualization system that allows clinicians to evaluate diagnostic imaging studies of CT and MR image data interactively in two or three dimensions. Proprietary clinical protocols automatically adapt the visualization settings to the examination type. Protocols are applied to each patient exam, producing up to six different 2D or 3D views. Vitrea is network ready and DICOM 3.0 conformant, and also has Internet authoring capabilities. Vitrea 2, a faster version due out soon, is compatible with Windows NT.

Reader Service No. 114

These notes are compiled in the Nature office from information provided by the manufacturers. For more details, fill in the reader service card.

Diode Pumped Solid State Laser

Q-Peak Inc. www.qpeak.com

Peak efficiency from latest 355 nm laser

The MPV-355 QS 10 diode pumped solid state laser offers more than 1W of 355 nm wavelength at 50 kHz pulse rate. It is the latest addition to Q-Peak's line up of IR, Green and UV wavelength diode pumped solid state laser systems, which incorporate Q-Peak's patented 'Gain Module' technology. This design, with high beam quality, effective use of diode pump power, scalability and inherent simplicity, allows for both high power generation and high harmonic conversion efficiency.

Reader Service No. 109

LTB compact nitrogen lasers

Photonic Solutions www.pspic.com

Pocket-sized nitrogen and dye laser systems

LTB compact nitrogen lasers are suitable for a wide range of applications including biological analysis, medicine and environmental protection. The PNL series, which runs from either battery or mains, is available in two versions: PNL100 has 100-kW peak power and pulse energies of >30µJ, and PNL200 has

ADVERTISEMENTS

Homogenizers?

- Tissue-Tearor™
- BeadBeater™
- Mini-BeadBeater™
- BioHomogenizer™

Check Us Out

www.biospec.com

BioSpec Products, Inc.
Ph/Fax 918-336-3363
e-mail info@biospec.com

New Probe Sizes
0.45cm!

**Long Read
DNA Sequencing**

Up To 1200 bp/read

Fast Turnaround Time

Only \$50/read

MacConnell Research Corp
1-800-466-7949

**Detect residual DNA with Wako's
DNA Extractor Kit
or Isolate Genomic DNA with
DNA Extractor WB Kit**

Both use the Nal Method
No Phenol No Chloroform

Wako

Wako BioProducts
Wako Chemicals USA, Inc.
1600 Bellwood Road
Richmond, VA 23237
Telephone: (804) 271-7677
Facsimile: (804) 271-7791
(800) 992-9256
bioproducts@wakousa.com